Near-field spectroscopy of semiconductor device structures and plasmonic crystals

D I S S E R T A T I O N

zur Erlangung des akademischen Grades
doctor rerum naturalium
(dr. rer. nat.)
im Fach Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

von
Herr Dipl.-Phys. ViktorMalyarchuk
geboren am 13.05.1973 in Ternopil (Ukraine)

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. J. Mlynek

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Dr. M. Linscheid

Gutachter:

1. Prof. Dr. Thomas Elsässer
2. Prof. Dr. Oliver Benson
3. Prof. Dr. Paul Fumagalli

eingereicht am: 25. April 2003
Tag der mündlichen Prüfung: 10. September 2003

Contents

Acknowledgments

I feel indebted to many people from our scientific team who made significant contribution into success of presented work.

At fist place I would like to thank Prof. T. Elsässer for opportunity to participate in the work of his group on delightsome projects involving bleeding edge near-field spectroscopy, and for his interest at all stages of investigations. I have greatly benefited from his leadership and experience.

I am deeply appreciating the knowledge that I acquire from Dr. J. W. Tomm. His contribution in understanding photocurrent near-field spectroscopy can hardly be overestimated. I thank him for productive discussions, critical readings of the manuscript and for the help I have been getting during my work at MBI.

I am grateful to Dr. Ch. Lienau for his dedication and uncompromised maintenance of high-quality scientific standards. I learn from him what is to be a real scientist. I thank him for supervision of our experiments, informative discussions and careful readings of the manuscript.

I am thankful to Prof. D. S. Kim for providing us with plasmonic samples, for time spent in laboratory and for stimulating discussions.

I owe to T. Günther, S. C. Hohng and Y. C. Yoon many days shared in the laboratory. I thank them for their contribution in success of experimental part of this work.

I am thankful to Dr. R. Müller for his FDTD calculations and for his patience and help in improving my German during discussions about plasmon physics.

I express my gratitude to Monika Tischer for patience and reliability in the fabrication of probes and for hours spent in REM-laboratory.

I want to thank Dr. V. Kalosha and Dr. A. Husakou for numerous theoretical tips and tricks.

I am thankful to members of the department (current and former) including Felix Eickemeyer, Francesca Intonti, Julian Edler, Matteo Rini, Jens Stenger, Valentina Emiliani, Nils Huse, Markus Raschke, Karsten Heyne, Kerstin Müller, Sendy Schwirzke-Schaaf, Axel Gerhardt, Thomas Unold for all kind of help I have always got when I needed it and for a nice friendly working atmosphere.

Particular thanks to friends for their support, understanding and patience for my permanent lack of time.

I am grateful to my parents for their continuous support in all my endeavors.

And very spacial thanks and admiration to my wife Iryna. Her love, patience, understanding and support helped me enormously to succeed in this challenge.

Publications in conjunction with this thesis

[1] T. Guenther, V. Malyarchuk, J. W. Tomm, R. Mueller, C. Lienau, and J. Luft, "Near-Field Photocurrent Imaging of the Optical Mode Profiles of Semiconductor Laser Diodes," Appl. Phys. Lett. **78,** 1463–1465 (2001).

[2] S. C. Hohng, V. Malyarchuk, C. Lienau, and D. S. Kim, "Surface plasmon optic devices and radiating surface plasmon sources for photolithography," Invention disclosure (2001), internal Application No. PCT/KR01/01921.

[3] A. Maaßdorf *et al.*, "Minority-carrier kinetics in heavily doped GaAs:C studied by transient photoluminescence," J. Appl. Phys. **91,** 5072–5078 (2002).

[4] J. W. Tomm *et al.*, "The impact of defects to minority-carrier dynamics in heavily doped GaAs:C analyzed by transient photoluminescence spectroscopy," Materials Science & Engineering B (Solid-State Materials for Advanced Technology) **B91-92,** 25–28 (2002).

[5] V. Malyarchuk, J. W. Tomm, C. Lienau, F. Rinner, and M. Baeumler, "Nanoscopic Measurements of Surface Recombination Velocity and Diffusion Length in a Semiconductor Quantum Well," Appl. Phys. Lett. **81,** 346–348 (2002).

[6] V. Malyarchuk, J. W. Tomm, C. Lienau, M. Behringer, and J. Luft, "Uniformity tests of individual segments of interband cascade diode laser Nanostacks®," J. Appl. Phys. **92,** 2729–2733 (2002).

[7] S. C. Hohng *et al.*, "Light emission from the shadows: Surface plasmon nano-optics at near and far field," Appl. Phys. Lett. **81,** 3239–3241 (2002).

[8] J. W. Tomm, A. Gerhardt, R. Müller, V. Malyarchuk, Y. Sainte-Marie, P. Galtier, J. Nagle, and J.-P. Landesman, "Spatially resolved spectroscopic strain measurements on high-power laser diode bars," J. Appl. Phys. **93,** 1354–1362 (2003).

[9] D. S. Kim, S. C. Hohng, V. Malyarchuk, Y. C. Yoon, Y. H. Ahn, K. J. Yee, J. Park, J. Kim, Q. H. Park, and C. Lienau, "Microscopic origin of surface-

plasmon radiation in plasmonic band-gap nanostructures," Phys. Rev. Lett. **91,** 143901(1–4) (2003).

[10] S. C. Hohng, Y. C. Yoon, D. S. Kim, V. Malyarchuck, C. Lienau, J. W. Park, and K. H. Yoo, "Near-field studies on light emission from a one-dimensional metal grating," Appl. Phys. Lett. (2003), submitted for publication.

[11] V. Malyarchuk, S. C. Hohng, J. W. Park, Y. C. Yoon, C. Lienau, K. H. Yoo, J. Kim, S. H. Han, Q.-H. Park, and D. S. Kim, "Evolution of the Near-Field Patterns into the Far-Field in Surface Plasmonic Band Gap Nano-Structures," Opt. Lett. (2003), submitted for publication.

Conference contributions in conjunction with this thesis

[1] V. Malyarchuk, J. W. Tomm, T. Gunther, and C. Lienau, "A novel contrast mechanism of emission mode near-field microscopy applied to diode laser structures," In *Summaries of papers presented at the Conference on Lasers and Electro-Optics*, Postconference Technical Digest (IEEE Cat. No. 01CH37170) p. 196 (USA, Washington, DC, USA, Opt. Soc. America, 2001, 2001).

[2] S. C. Hohng, Y. C. Yoon, D. S. Kim, V. Malyarchuck, C. Lienau, J. W. Park, K. H. Yoo, and H. Y. Ryu, "Transmission, control and coherent propagation of surface plasmons in metal nanostructures," In *Summaries of papers presented at the Quantum Electronics and Laser Science Conference*, Postconference Technical Digest (IEEE Cat. No.01CH37172) pp. QPD4–1–2 (USA, Washington, DC, USA, Opt. Soc. America, 2001, 2001).

[3] V. Malyarchuck, J. W. Tomm, T. Günther, R. Müller, R. Kunkel, C. Lienau, and J. Luft, "Large Optical Cavity waveguides for high-power diode laser applications," In *Proceedings of the SPIE – The International Society for Optical Engineering*, **4287,** 111–117 (San Jose, CA, USA, 2001).

[4] C. Lienau, T. Guenther, V. Malyarchuk, and J. W. Tomm, "Analysis of waveguide architectures of high-power diode lasers by near-field scanning optical microscopy," In *CLEO/Pacific Rim 2001. 4th Pacific Rim Conference on Lasers and Electro-Optics*, Technical Digest (IEEE Cat. No. 01TH8557) **1,** I–180–1 (Japan, Piscataway, NJ, USA, IEEE, 2001, 2001).

Curriculum Vitae

Surname:	Malyarchuk
Fist name:	Viktor
Date of Birth:	May 13th, 1973
Place of Birth:	Ternopil, the Ukraine
Nationality:	Ukrainian
Marital Status:	Married since January 10th, 2001
1980-1990	School №15, Ternopil, the Ukraine
1990-1995	Physical department of "Taras Schevchenko" National University, Kiev, the Ukraine Master thesis: *"Photoluminescence of the filament-like crystals of silicon"* Thesis adviser: Prof. M. Ya. Valakh
1995-1999	scientific co-worker in Institute of Semiconductor Physics National Academy of Science of Ukraine
1999-2003	Ph.D. student in the group of Prof. T. Elsässer, Max-Born-Institute, Berlin, Germany Ph.D. Thesis: *"Near-field spectroscopy of semiconductor device structures and plasmonic crystals"* Thesis Advisor: Prof. T. Elsässer

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbstständig angefertigt und keine weitetern als die angegeben Hilfsmittel verwendet zu haben.

Berlin, den 24. April 2003 Viktor Malyarchuk

Zusammenfassung

Die Abmessungen der aktiven Gebiete moderner optoelektronischer Bauelemente bewegen sich typischerweise in Regionen, die der Lichtwellenlänge vergleichbar oder noch kleiner sind. Die überwiegende Mehrzahl der optischen Phänomene, die in solchen Bauelementen auftreten können aber durchaus mit klassischen Ansätzen, die auf den Maxwell-Gleichungen beruhen, verstanden werden. Trotzdem es sich also unter diesem Aspekt um durchaus „klassische Systeme“ handelt, ist ihre Erforschung mit klassischer Optik, die den Grenzen des Beugungslimits unterliegt, nur unvollständig möglich.

In dieser Arbeit diskutieren wir die Physik zweier solcher Systeme, nämlich die von Halbleiterbauelementestrukturen und die plasmonischer Kristalle.

Der Laser zählt zu den Schlüsseltechnologien in hoch entwickelten Gemeinwesen. Die Nutzung von Lasern zur Materialbearbeitung beschleunigt den Produktionsprozess bezüglich Geschwindigkeit, Qualität, Zuverlässigkeit und Flexibilität. Sowohl bezüglich des technischen Potentials als auch der Markdurchdringung ist die Lasertechnik von einer Sättigung immer noch weit entfernt. Eine der Ursachen besteht darin, dass Produktionsprozesse noch nicht bezüglich des Lasereinsatzes adaptiert sind. Weiterhin ist festzustellen, dass heutige Lasersysteme noch relativ schwer und voluminös sind und auch erhebliche Kosten für Betrieb und Wartung anfallen.

Die Situation hat sich allerdings mit dem Auftauchen der Diodenlaser in einer Weise verändert, wie es zuvor mit der Elektronik geschah, als der Transistor auftauchte. Heutzutage sind Hochleistungslaserdioden Schlüsselelemente vieler Lasersysteme und stehen im Wettbewerb mit Gaslasern und lampengepumpten Festkörperlasersystemen. Derartige Systeme finden immer mehr den Weg zu industriellen Anwendungen, da sie kompakt sind, einen Wirkungsgrad von über 50% haben und nur einen geringen Kühlaufwand erfordern. Bezüglich des Wirkungsgrades liegt damit eine etwa fünffache Überlegenheit gegenüber konventionellen Systemen vor. Ein weiterer Grund für den Vormarsch der Diodenlaser ist darin zu suchen, dass die Kosten für eine bestimmte Diodenlaserleitung [€/W] immer mehr fallen.

Zweidimensionale plasmonische Kristalle haben ein großes Potential für praktische Anwendungen für künftige optoelektronische Bauelemente. Ihre optischen Eigenschaften werden wesentlich durch evaneszente Oberflächenplasmon-Moden bestimmt. Diese ermöglichen, durch Nutzung von Interferenzeffekten, optische Anregungen auf einer sub-Wellenlängenskala zu lokalisieren. Damit ergeben sich

neue Möglichkeiten sowohl zur Lichtlokalisierung als auch zum Transport optischer Anregungen auf Nanometer-Längenskalen.

Plasmonische Gitter sind auch vom Standpunkt der Physik einzigartig. Um ein festkörperphysikalisches System vollständig zu beschreiben, bedarf man Informationen über die Quasiimpulsdispersion (Brillouinsche Zone), die Eigenmoden und die Verlustmechanismen, die im System wirksam sind. Üblicherweise ist es durchaus möglich Informationen über die Brillounsche Zone und die Verlustmechanismen zu erhalten, während sich die Eigenmoden allerdings sehr häufig dem detaillierten experimentellen Zugriff entziehen. Eine einzigartige Eigenschaft plasmonischer Kristalle besteht darin, dass man direkt experimentell Informationen über alle drei genannten Eigenschaften erzielen kann. Das macht sie auch als Modellsysteme für die Grundlagenphysik interessant.

Die Nahfeldmikroskopie erlaubt es, die durch das Beugungslimit für die klassische Mikroskopie vorgegeben Grenzen zu überschreiten. Die von Synge in den zwanziger Jahren des letzten Jahrhunderts geäußerte Grundidee besteht darin, die oberflächennahen, sich nicht ausbreitenden Lichtwellen, also das „optische Nahfeld“, zur Bilderzeugung zu nutzen. Aus Mangel an technischen Möglichkeiten verzögerte sich die Realisierung dann allerdings bis zu einem Zeitpunkt, der jetzt etwa 20 Jahre zurückliegt. Die Auflösungsgrenze konventioneller optischer Systeme hängt von der numerischen Apertur des genutzten Linsensystems und der betrachteten Lichtwellenlänge ab. Im Gegensatz dazu wird die Auflösungsgrenze eines Nahfeldmikroskopes ausschließlich durch die Abmessungen der Öffnung der Nahfeldsonde bestimmt. Die Sonde kann sowohl zur Erzeugung eines Nahfeldes als auch zur Abfrage des Nahfeldes an einer Oberfläche eingesetzt werden und möglicherweise auch beide Funktionen gleichzeitig ausüben. Übliche Sonden sind gezogene oder geätzte Glasfaserspitzen, die mit einem Metallüberzug versehen sind. An der äußersten Spitze befindet sich eine Öffnung in der Metallisierung, die typischerweise einen Durchmesser im nm-Bereich hat. Führt man diese Öffnung nun systematisch in geringem Abstand über eine Probenoberfläche, kann man die gemessenen Signale zur Bilderzeugung einsetzen. Damit gehört die Nahfeldmikroskopie zur Klasse der Rastermikroskope.

Ein entscheidender Vorteil der Nahfeldmikroskopie besteht nun in der Möglichkeit, die extrem hohe räumliche Auflösung (oft besser als 100 nm) mit anderen spektroskopischen Techniken zu kombinieren. So gibt es verschiedenste Kontrasterzeugungsmechanismen, die zur Bildgebung genutzt werden können, wie die lokale Absorption, Lumineszenz, Reflexion oder auch Polarisation.

Das Modenprofil von Wellenleitern in optoelektronischen Bauelementen, die typischerweise Abmessungen im sub-μm-Bereich haben, ist mit konventioneller optischer Spektroskopie wegen der Beugungsbegrenzung nicht erforschbar. Hier demonstrieren wir das Potential der optischen Nahfeldmikroskopie beim Überwinden der Beugungsbegrenzung. Unser Mikroskop kombiniert die hohe Ortsauflösung mit der Möglichkeit der selektiven Anregung mittels durchstimmbaren Laser, was die gezielte Erforschung des Wellenleiters der Bauelementestruktur ermöglicht.

Experimente an InGaAs/AlGaAs-Hochleistungslaserdiodenarrays mit unterschiedlichen Wellenleiterarchitekturen zeigen direkt die Auswirkungen unterschiedlicher Wellenleiter auf

- die Anzahl der im Wellenleiter geführten Moden und
- den räumlichen Verlauf des Profils der Grundmode und von Moden höherer Ordnung.

Experimente mit monolithisch übereinander gewachsenen Diodenstrukturen, in denen es zwei nominell identische übereinander angeordnete Wellenleitersegmente, die durch einen Tunnelübergang voneinander separiert sind, gibt, demonstrieren, dass optische Nahfeldmikroskopie

- einen direkten und separaten Zugriff zu beiden optisch aktiven Gebieten erlaubt,
- die Analyse der Bauelementeemission, speziell der Elektrolumineszenz und der Laseremission, sowie auch der Photolumineszenz und des Photostromes ermöglicht und auf der Basis dieser Daten eine konsistente Diskussion der Bauelementeeigenschaften erfolgen kann.
- Es wird demonstriert, dass eine verringerte Laseremission eines Segmentes nicht etwa thermisch bedingt ist, sondern auf eine erhöhte Haftstellenkonzentration in oder in der Umgebung der Quantentröge im Wellenleiter zurückgeführt werden kann.
- Nahfeldmikroskopie erlaubt ebenfalls den Vergleich der Potentialgradienten, die in den einzelnen Wellenleitersegmenten vorliegen.

An der Spaltfläche einer Quantentrogstruktur modifiziert die Oberflächenrekombination das Photolumineszenzsignal auf einer räumlichen Skale von einigen μm in nachhaltiger Weise. Die notwendige räumliche Auflösung zur Erforschung dieses Phänomens kann durch optische Nahfeldmikroskopie bereitgestellt werden. Aus derartigen Experimenten konnten wir die nachfolgend aufgeführten Informationen extrahieren:

- Der räumliche Verlauf des Photolumineszenzsignal an einer Probenkante wird stark durch die gewählte optische Anregungsdichte beeinflusst.
- Der Einfluss der Oberflächenrekombination auf das Photolumineszenzsignal nimmt mit zunehmender Anregungsdichte ab, was mit einer Sättigung nichtstrahlender Rekombinationszentren erklärt wird.
- Aus dem Vergleich der Daten mit der Lösung einer zweidimensionalen Bilanzgleichung können wir unabhängig Oberflächenrekombination und Diffusionslänge bestimmen.

Wenn Licht auf einen vergleichsweise dicken Metallfilm fällt, der mit sehr kleinen Löchern versehen ist, entstehen auf der Rückseite Schatten. Wir zeigen, dass

- auf der Nanoskala auch Licht von der Rückseite emittiert werden kann. Dabei kann die Intensität der Emission von der Metallisierung zwischen den Löchern größer sein als unmittelbar an den beleuchteten Öffnungen;
- die Symmetrie der Nahfeldemissionsverteilung durch die Symmetrieeigenschaften der Oberflächenplasmonwellen bestimmt wird;
- die Nahfeldemissionsverteilung soweit im Fernfeld erhalten bleiben kann, bis die Verteilung einer Sinusfunktion entspricht;
- die sehr unübliche Lichtemission aus den abgeschatteten Bereichen durch strikte Wellenvektorauswahlregeln bestimmt wird;
- die Ausbreitungslänge und Zerfallszeit der Oberflächenplasmonanregungen in plasmonischen Kristallen direkt gemessen werden können.
- Rayleigh Streuung
  - wird als der mikroskopische strahlende Streumechanismus der Plasmonanregungen an den Nano-Löchern identifiziert und
  - erweist sich gleichfalls als die Ursache einer endlichen Linienbreite der in Transmission beobachteten Resonanzen.
- Eine erste Bandstrukturanalyse der Transmissionsspektren der plasmonischen Kristalle zeigt klar die Existenz einen Bandlücke in der Bandstruktur.

Wir glauben, dass die wissenschaftlichen Ergebnisse, die im Rahmen dieser Arbeit erzielt wurden, klar zeigen, das

- die optische Nahfeldmikroskopie eine sensitives Werkzeug für nichtzerstörenden Analytik an optoelektronischen Bauelementen und plasmonischen Kristallen ist.
- Weiterhin beleuchten unsere Ergebnisse einen neuen Aspekt der Eigenschaften plasmonischer Kristalle. Die Klärung der Ursachen der verstärkten Transmission erlaubt neue mikroskopische Einsichten zur Physik der Licht–Materie–Wechselwirkung in plasmonischen Nanostrukturen und liefert bedeutende neue Informationen darüber, wie man die Eigenschaften optoelektronischer Bauelemente und plasmonischer Nanostrukturen noch besser maßschneidern kann.

Die Arbeit ist wie folgt strukturiert:

Kapitel 1 beinhaltet diese Einführung.

Kapitel 2 ist der theoretischen Basis der Elektrodynamik gewidmet. Die Eigenschaften der Maxwell-Gleichungen, das elektromagnetische Eigenwertproblem und die Ursachen für das Auftreten der Beugungsbegrenzung werden betrachtet. Die theoretischen Konzepte, die der optischen Nahfeldspektroskopie zu Grunde liegen, werden ebenfalls eingeführt.

Kapitel 3 beschreibt die für die Arbeiten benutzte Messapparatur. Wir diskutieren die Eigenschaften der Fasersonde, die als Schlüsselelement des optischen Nahfeldmikroskopes fungiert. Der Rückkopplungsmechanismus und seine Realisierung in unserer Apparatur werden diskutiert. Das gesamte Zimmertemperaturnahfeldmikroskop, die Anregungslichtquellen und auch die Nachweismethoden werden beschrieben.

Kapitel 4 befasst sich mit spektroskopischen Untersuchungen an Halbleiterbauelementestrukturen. Züchtungsmethoden, die physikalischen Grundlagen der Prozesse, die in Halbleiterstrukturen ablaufen und die Hauptelemente von Halbleiterdiodenlasern werden betrachtet. Die Methode des „virtuellen Wellenleiters“ zur Eigenmodenberechnung von Wellenleitern wird vorgestellt. Ebenso präsenieren wir Photostromdaten von Diodenlasern mit einem und auch zwei übereinander befindlichen Wellenleitern. Eine neue Methode zur simultanen Bestimmung von Oberflächenrekombinationsgeschwindigkeit und Diffussionslänge wird vorgestellt.

Kapitel 5 beschäftigt sich mit Plasmon-Nanooptik. Wir beginnen mit einer Übersicht der relevanten Besonderheiten von Metallen, den Grundlagen der Oberflächenplasmonphysik und den theoretischen Methoden zur Simulation der Eigenschaften photonischer und plasmonischer Kristalle. Danach präsentieren wir experimentelle und theoretische Ergebnisse zur Feldverteilung an der Oberfläche von Nano-Loch-Arrays. Experimentelle Ergebnisse, die in der Zeitdomäne erzielt wurden, Ergebnisse zur Bandstrukturanalyse sowie Ergebnisse für Arrays mit unterschiedlichen Nano-Loch-Abmessungen erlauben die Aufklärung der Streumechanismen sowie der mikroskopischen Quellen der Oberflächenplasmonemission. Wir zeigen Ergebnisse zur Polarisations- und Wellenlängenabhängigkeit der Nahfeldemissionsverteilung von Oberflächenplasmonen. Im letzten Teil dieses Kapitels beschäftigen wir uns mit der experimentellen und theoretischen Analyse des Überganges von der Nahfeld- zur Fernfeldverteilung.

Kapitel 6 beinhaltet die Zusammenfassung der erzielten Ergebnisse.

Chapter 1

Introduction

The dimensions of most modern optoelectronic devices belong to the size region that is comparable or smaller than the light wavelength. The vast majority of optical phenomena that take place in such devices can be understood with help of classical considerations based on Maxwell's equations. Despite being classical systems experimental investigations by means of conventional optics hardly can bring full information about such devices because of fundamental limitations imposed by the diffraction limit.

In this work we will discuss two types of such systems, namely semiconductor device structures and plasmonic crystals.

The laser is one key technology in highly industrialized economies. Using of lasers in material processing brings significant boost for fabrication process in speed, quality, reliability, and flexibility. With respect to its technical potential and market diffusion, laser processing is still far away from saturation. One of the reasons for this is due to the fact that process technologies are not sufficiently adapted to laser employment. Furthermore, even today's laser system are relatively heavy and voluminous with relatively high costs for operation and maintenance.

The situation changed fundamentally with the advent of diode lasers very much in the way as the transistor revolutionized electrical engineering. Nowadays high power semiconductor lasers are key elements of a new breed of laser systems that are competing with gas lasers and lamp pumped solid state lasers. Such devices increasingly find their way into industrial applications, as they are compact, easy to cool, deliver power efficiency beyond 50%, which is about five times higher than any other kind of conventional laser has to offer. One more reason for this success is that the production costs for diode lasers are becoming more and more reduced.

Two-dimensional plasmonic crystals have a big potential for practical uses in future optoelectronic devices. Their optical properties are governed by evanescent surface plasmon excitations. These allow, by using interference effects, to localize optical excitations on a sub-wavelength length scales. This offers new possibilities both for light localisation and for the transport of optical excitations on a nanometer length scale.

They are also unique from a fundamental physics point of view. In order to fully characterize the physical system in solid state physics it is necessary to get knowledge about the systems dispersion relation (Brillouin zone), shape of eigenmodes, and information about losses in the system. Usually it is possible to get experimental information about the Brillouin zone and losses but it appears mostly impossible obtaining details about the eigenmodes. The uniqueness of the plasmonic crystal is that one can get access to all three types of information directly. This make them extremely useful as a model system for understanding the fundamental physics of periodic media.

Near-field scanning optical microscope (NSOM) overcomes the principal diffraction limit of the conventional light microscopy. The fundamental idea of using light localization in nanoaperture for picture formation was proposed by Synge in the twenties of last century. The lack of appropriate technologies at that time delayed the practical realization of such a technique until two decades ago. The resolution limit of a conventional optical system depends on the numerical aperture of the employed lens system and the used wavelength. In contrast, the resolution of the NSOM depends exclusively on near-field probe size. The probe serves for near-field creation and/or near-field detection of the investigated structure. Commonly used probes are made from pulled/etched glass-fiber tips that is coated with metal. At the very end of it an aperture is situated that is only a few nanometers in diameter. A two-dimensional picture of the object is produced by raster scanning the tip over the sample surface. So near-field microscopy is a typical example of raster scanning microscopy.

The biggest advantage of the near-field microscopy is the possibility to combine an extremely high lateral resolution (less than 100 nm) with other techniques of optical spectroscopy. There are many contrast mechanisms that can be exploited for optical near-field investigations, namely, the local absorption and luminescence spectroscopies as well as space resolved reflection and polarization spectroscopies.

The mode profile of submicron-sized waveguides in modern optoelectronic devices can not be mapped by means of conventional optics due to the diffraction limit. Here we demonstrate the potential of near-field photocurrent spectroscopy for breaking such limitations. The technique combines the submicron spatial resolution of near-field optics with tunable laser excitation, allowing for selective investigation of the waveguide properties of the device structure. Experiments on InGaAs/AlGaAs high-power laser diodes with different waveguide designs provide direct visualization of the effect of the waveguide design on:

- the number of guided modes and
- the spatial profile of both fundamental and higher-order modes.

From investigations of monolithically stacked diode lasers with two nominally identical waveguide segments separated by a tunnel junction we see that:

- near-field optical microscopy provides straightforward and separate access to the properties of both optically active segments;
- device emission, namely electroluminescence and lasing, as well as photoluminescence and photocurrent data, can be recorded with high spatial resolution and consistently interpreted;
- reduced laser emission from the laser segment that is situated closer to the substrate is not caused by thermal effects but due to a larger trap concentration within or in the vicinity of the quantum wells of this laser segment;
- in the unbiased devices the potential gradient in the segment with reduced laser emission is significantly larger than in the one closer to the heat sink.

Near the edge area of a quantum well, surface recombination gives rise to a gradual variation of the photoluminescence signal on a micrometer length scale. Reliable spatial resolution for the photoluminescence measurement in this area can be provided through the near-field microscopy. From such experiments we are able to extract following information:

- the overall shape of the luminescence signal in this transition region depends strongly on the excitation intensity.
- the surface recombination velocity decreases with increasing intensity due to the saturation of nonradiative defect states;
- from solving two dimensional diffusion equations for the obtained data we extract the surface recombination velocity and the diffusion length;

When light illuminates a thick metal film perforated with small holes, shadows appear. We find that:

- at the nanoscopic level, light can be emitted from the back side predominantly from the metal surfaces between the holes–shadows can be brighter than the lighted holes;
- the symmetry of the near-field emission pattern is determined by the symmetry of the surface plasmon waves;
- nanoscopic emission patterns from the metal can be preserved to the far-field region, where the pattern becomes sinusoidal;
- this unusual behavior of light emission from the shadows is explained by efficient wave vector selection;
- the propagation length and damping time of SP excitations in plasmonic crystals can be measured directly;
- Rayleigh scattering is:

  - the microscopic mechanism responsible for radiative damping of surface plasmon excitations at the nano-holes;
  - the origin of the finite linewidth of the transmission resonances;
- the first lineshape analysis of the transmission spectra of plasmonic crystal prove the existence of a band gap in such a system.

We believe that scientific results produced in the scope of this work

- justify that near-field microscopy is a sensitive nanoscopic tool for nondestructive analysis of optoelectronic and plasmonic devices.
- highlight an interesting new aspect of the origin of enhanced transmission, gives new insight into the microscopic physics of light–matter–interaction in plasmonic nanostructures, and provides important new information on how to tailor the optical properties of optoelectronic devices and plasmonic nanostructures.

Thesis is organized as follows:

Chapter 1 is this introduction.

Chapter 2 is devoted to theoretical basis of electromagnetic phenomena. Properties of Maxwell's equations, electromagnetic eigenvalue problem, and the reasons for diffraction limit are considered. Theoretical concepts of near-field scanning optical microscopy are introduced.

Chapter 3 describes the experimental technique used. We discuss the fiber tip as a key element of a near-field scanning optical microscope. The feedback mechanism and its particular realization in our setup are discussed. The room temperature near-field microscope as well as excitation sources and detection methods are described.

Chapter 4 deals with spectroscopy of semiconductor device structures. Device growth methods, basic physical processes that take place in semiconductor, basic elements and properties of semiconductor diode laser are considered. The method of the fake waveguide for eigenmode calculations as well as near-field photocurrent imaging experiments on single waveguide semiconductor lasers and on novel monolithic diode laser stacks are presented. A novel technique for the determination of the surface recombination velocity and diffusion length is introduced.

Chapter 5 treats plasmon nano-optics. We start from an overview on particularities of metals, basic properties of surface plasmons, and the methods used for theoretical simulation of photonic and plasmonic crystals. Then we present experimental and theoretical results highlighting the field distribution on the surface of nano-hole arrays. The time domain experiments and

band gap analysis as well as hole size dependence experiment for surface plasmon scattering clarify the macroscopic origin of surface plasmon radiation in such systems. We present polarization dependence and wavelength dependence of surface plasmons near-field emission patterns obtained from polarization controlled experiment. The last topics of this chapter are experimental and theoretical analysis that deal with evolution of the near-field patterns into a far-field.

Chapter 6 is the summary of the thesis.

Chapter 2

Introduction into nano-optics

In this chapter we will provide the theoretical basis required for further discussions addressing semiconductor device structures and plasmonic crystals. The finite spatial resolution limitations of the conventional far-field optics will be shown and ways how to overcome them are discussed.

In the beginning we will consider Maxwell's equations as the basis of electromagnetic theory. We will discuss the main properties of electromagnetism and conditions when the classical approach can be used. We will consider wave equations for macroscopic media and will show that in a plane wave approach we can treat those equations as eigenvalue problem (both the general eigenvalue problem and the electromagnetic eigenvalue problem than will be described). We will address the scaling properties of Maxwells equations and will make comparison with quantum mechanics.

Due to the sub-wavelength scale of the objects under investigation, diffraction and complications, that it brings, play an important role. We will consider some aspects of the Kirchhoff diffraction theory. Origins of the diffraction limit in the far field of point like source, methods of it's overcoming and shortcomings in this methods will be discussed.

We will address the electromagnetic fields generated by a Hertzian dipole. The propagating part of the solution responsible for far-field (FF) radiation and the evanescent near-field (NF) part will be described. After the description of the concepts of the near-field technique we consider criteria that define the resolution limit of the NSOM. With this knowledge we will look at commonly used experimental NSOM configurations.

2.1 Maxwell equations

2.1.1 Maxwell equations and general assumptions

Maxwell equations are the equations that governing classical electromagnetic phenomena (1):

$$\nabla \cdot \mathbf{D} = 4\pi\rho \tag{2.1}$$

$$\nabla \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c}\mathbf{J} \tag{2.2}$$

$$\nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = 0 \tag{2.3}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{2.4}$$

where for a external source in vacuum, $\mathbf{D} = \varepsilon_0 \mathbf{E}$ and $\mathbf{B} = \mu_0 \mathbf{H}$. The first two equations then become

$$\nabla \cdot \mathbf{E} = 4\pi\rho/\varepsilon_0 \tag{2.5}$$

$$\nabla \times \mathbf{B} - \frac{\partial \mathbf{E}}{c^2 \partial t} = 4\pi\mu_0 \mathbf{J} \tag{2.6}$$

Implicit in the Maxwell equations is the continuity equation for the charge density ρ and the current density $\mathbf{J}$,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \tag{2.7}$$

This follows from combining the time derivative of the first equation (2.1) with the divergence of the second equation (2.2). Also essential for the consideration of charged particle motion is the Lorentz force equation,

$$\mathbf{F} = q(\mathbf{E} + v \times \mathbf{B}) \tag{2.8}$$

which gives the force acting on a point charge q in the presence of electromagnetic fields.

There is a lack of symmetry in the appearance of the source terms in Maxwell's equations. The first two equations (2.1, 2.2) have sources; the second two (2.3, 2.4) do not. This reflects the experimental *absence of magnetic charges* and currents.

2.1.2 Wave equation

Most of the phenomena considered in this work originate from electromagnetic wave propagation in macroscopic media. Such propagating waves can be reliably described through wave equation that derived from Maxwell equations and inherit some particularities form surrounding media. Here and in few following subchapters we consider the simplest case: the propagation of electromagnetic waves through a dielectric medium with no free charges or currents (2). In such a medium with no sources of light, we set $\rho = \mathbf{J} = 0$.

A few standard assumptions are necessary. First, we assume that the field strength is small enough, that we are in the linear regime. The second assumption is that the material is macroscopic and isotropic, so that the dielectric constant is a scalar. Third, we ignore any explicit frequency dependence of the dielectric constant. Fourth, we focus only on low-loss materials.

Taking into account all this assumptions and expanding the fields into a set of harmonic modes:

$$\mathbf{H}(\mathbf{r},t) = \mathbf{H}(\mathbf{r})e^{i\omega t} \tag{2.9}$$

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}(\mathbf{r})e^{i\omega t} \tag{2.10}$$

the two curl equations will take a form:

$$\nabla \times \mathbf{E}(\mathbf{r}) + \frac{i\omega}{c}\mathbf{H}(\mathbf{r}) = 0 \tag{2.11}$$

$$\nabla \times \mathbf{H}(\mathbf{r}) + \frac{i\omega}{c}\varepsilon(\mathbf{r})\mathbf{E}(\mathbf{r}) = 0 \tag{2.12}$$

We can decouple these equations in the following way. Divide the (2.12) equation by $\varepsilon(\mathbf{r})$, and then take the curl. Then use the (2.11) equation to eliminate $\mathbf{E}(\mathbf{r})$. The result is an equation entirely in $\mathbf{H}(\mathbf{r})$:

$$\nabla \times \left(\frac{1}{\varepsilon(\mathbf{r})}\nabla \times \mathbf{H}(\mathbf{r})\right) = \left(\frac{\omega}{c}\right)^2 \mathbf{H}(\mathbf{r}) \tag{2.13}$$

This is the *wave equation*. In addition to the conditions $\nabla \cdot \mathbf{H}(\mathbf{r}) = \nabla \cdot \mathbf{D}(\mathbf{r}) = 0$ that follows from divergence equations (2.1) and (2.4), it completely determines $\mathbf{H}(\mathbf{r})$. It is convenient to solve the wave equation to find the modes for $\mathbf{H}(\mathbf{r})$ for a given frequency, subject to the transversality requirement. Then use Eq. (2.12) to recover $\mathbf{E}(\mathbf{r})$:

$$\mathbf{E}(\mathbf{r}) = \left(\frac{-ic}{\omega\varepsilon(\mathbf{r})}\right)\nabla \times \mathbf{H}(\mathbf{r}) \tag{2.14}$$

2.1.3 Scaling properties of the Maxwell equations

One interesting feature associated to the electromagnetism in dielectric media is that there is no fundamental length scale other than the assumption that the system is macroscopic. There is no fundamental constant dimensions of length, so there is a simple relationship between the electromagnetic problems that differ only by a contraction or expansion of all distances (2).

For example, we have an electromagnetic eigenmode $\mathbf{H}(\mathbf{r})$ of frequency ω in a dielectric configuration $\varepsilon(\mathbf{r})$. The wave equation for this configuration is:

$$\nabla \times \left(\frac{1}{\varepsilon(\mathbf{r})}\nabla \times \mathbf{H}(\mathbf{r})\right) = \left(\frac{\omega}{c}\right)^2 \mathbf{H}(\mathbf{r}) \tag{2.15}$$

We now discuss the harmonic modes in a configuration of dielectric $\varepsilon'(\mathbf{r})$ that is just a compressed or expanded version of $\varepsilon(\mathbf{r})$: $\varepsilon'(\mathbf{r}) = \varepsilon(\mathbf{r}/\mathbf{s})$ for some scale parameter s. We can make a change of variables in (2.15), by using $\mathbf{r}' = s\mathbf{r}$ and $\nabla' = \nabla/s$:

$$s\nabla' \times \left(\frac{1}{\varepsilon(\mathbf{r}'/s)}\nabla' \times \mathbf{H}(\mathbf{r}'/s)\right) = \left(\frac{\omega}{c}\right)^2 \mathbf{H}(\mathbf{r}'/s) \tag{2.16}$$

$\varepsilon(\mathbf{r}'/s)$ is $\varepsilon'(\mathbf{r}')$. Dividing (2.16) by s results in

$$\nabla' \times \left(\frac{1}{\varepsilon'(\mathbf{r}')}\nabla' \times \mathbf{H}(\mathbf{r}'/s)\right) = \left(\frac{\omega}{cs}\right)^2 \mathbf{H}(\mathbf{r}'/s) \tag{2.17}$$

This is the wave equation again, with mode profile $\mathbf{H}'(\mathbf{r}') = \mathbf{H}(\mathbf{r}'/s)$ and the frequency $\omega' = \omega/s$. In words, if we want to know the new mode profile after changing the length scale by a factor s, we just scale the old mode and its frequency by the same factor. The solution of the problem at one length scale determines the solutions at all other length scales.

This fact is of considerable practical importance. For example, the microfabrication of complex micron-scale photonic crystals can be quite difficult. But models can be easily made and tested in the microwave regime, at the much larger scale of centimeters (3; 4). The model will have the same electromagnetic properties.

Just as there is no fundamental *length* scale, there is also no fundamental value of the dielectric constant. Suppose we know the harmonic modes of a system with dielectric configuration $\varepsilon(\mathbf{r})$, and we are curious about the modes of a system with a dielectric configuration that differs by a constant factor everywhere, so that $\varepsilon'(\mathbf{r}) = \varepsilon(\mathbf{r}/s^2)$. Substituting $s^2\varepsilon'(\mathbf{r})$ for $\varepsilon(\mathbf{r})$ in (2.15) yields

$$\nabla \times \left(\frac{1}{\varepsilon'(\mathbf{r})}\nabla \times \mathbf{H}(\mathbf{r})\right) = \left(\frac{s\omega}{c}\right)^2 \mathbf{H}(\mathbf{r}) \tag{2.18}$$

The harmonic modes of the new system are unchanged, but the frequencies are all scaled by a factor s: $\omega \to \omega' = s\omega$ If we multiply the dielectric constant everywhere by a factor of $1/4$, the mode patterns are unchanged but their frequencies double.

2.2 Eigenvalue problem

2.2.1 General eigenvalue problem

An *eigenvalue problem* is a typical situation that often arises in mathematical physics. It can by formulated as follows: we perform a series of operations, expressed by an operator on a function $\mathbf{F}(\mathbf{r})$ and if the result of a such action is just the function $\mathbf{F}(\mathbf{r})$ itself, multiplied by some constant, then we have a eigenvalue problem. This allowed function is called an *eigenfunction* or *eigenvector* of the operator, and the multiplicative constant is called the *eigenvalue*.

One restriction to the operator is: it should be a *Hermitian*. Let's show what it means for an operator to be Hermitian. First, in analogy with the inner product of two wave functions, we define the inner product of two vector fields $\mathbf{F}(\mathbf{r})$ and $\mathbf{G}(\mathbf{r})$ as

$$(\mathbf{F}, \mathbf{G}) \equiv \int d\mathbf{r}\mathbf{F}^*(\mathbf{r}) \cdot \mathbf{G}(\mathbf{r}) \tag{2.19}$$

Note that a simple consequence of this definition is that $(\mathbf{F}, \mathbf{G}) = (\mathbf{G}, \mathbf{F})^*$ for any $\mathbf{F}$ and $\mathbf{G}$. Also note that $(\mathbf{F}, \mathbf{F})$ is always real, even if $\mathbf{F}$ itself is complex. In fact, if $\mathbf{F}(\mathbf{r})$ is a harmonic mode of our electromagnetic system, we can always set $(\mathbf{F}, \mathbf{F}) = 1$ by using our freedom to scale any mode by an overall multiplier. Given $\mathbf{F}'(\mathbf{r})$ with $(\mathbf{F}', \mathbf{F}') \neq 1$ create

$$\mathbf{F}(\mathbf{r}) = \frac{\mathbf{F}'(\mathbf{r})}{\sqrt{(\mathbf{F}', \mathbf{F}')}} \tag{2.20}$$

From our previous discussion, $\mathbf{F}(\mathbf{r})$ is really the same mode as $\mathbf{F}'(\mathbf{r})$, since it differs only by an overall multiplier, but now we have $(\mathbf{F}, \mathbf{F}) = 1$. We say that $\mathbf{F}(\mathbf{r})$ has been *normalized*. Normalized modes is very useful in formal arguments. If however, one is interested in the physical energy of the field and not just its spatial profile, the overall multiplier is important.

Next, we say that an operator Ξ is *Hermitian* if $(\mathbf{F}, \Xi\mathbf{G}) = (\Xi\mathbf{F}, \mathbf{G})$ for any vector fields $\mathbf{F}(\mathbf{r})$ and $\mathbf{G}(\mathbf{r})$. That is, it does not matter which function is operated upon before taking the inner product. Clearly not all operators are Hermitian, but it is a case for electromagnetism.

2.2.2 Electromagnetic eigenvalue problem

The core of the electromagnetic problem for a harmonic mode in a dielectric medium is a wave equation for $\mathbf{H}(\mathbf{r})$ (2.13). The content of the equation is this: perform a series of operations on a function $\mathbf{H}(\mathbf{r})$, and if $\mathbf{H}(\mathbf{r})$ is an allowable electromagnetic mode, the result will just be a constant times the original function $\mathbf{H}(\mathbf{r})$. But this is a typical *eigenvalue problem*.

We identify the left side of the wave equation as an operator Θ acting on $\mathbf{H}(\mathbf{r})$:

$$\Theta\mathbf{H}(\mathbf{r}) = \left(\frac{\omega}{c}\right)^2 \mathbf{H}(\mathbf{r}) \tag{2.21}$$

Θ is the differential operator that takes the curl, than divides by $\varepsilon(\mathbf{r})$, and then takes the curl again:

$$\Theta\mathbf{H}(\mathbf{r}) \equiv \nabla \times \left(\frac{1}{\varepsilon(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r})\right) \tag{2.22}$$

The eigenvector $\mathbf{H}(\mathbf{r})$ are the field patterns of the harmonic modes, and the eigenvalues $(\omega/c)^2$ are proportional to the squared frequencies of those modes. One important thing to notice is that the operator Θ is a *linear* operator. That is, any linear combination of solutions is itself a solution; if $\mathbf{H}_1(\mathbf{r})$ and $\mathbf{H}_2(\mathbf{r})$ are both solutions of (2.21) with the same frequency ω, then so is $\alpha\mathbf{H}_1(\mathbf{r}) + \beta\mathbf{H}_2(\mathbf{r})$, where α and β are constants. For example, given a certain mode profile, we can construct another legitimate mode profile with the same frequency by simply doubling the field strength everywhere ($\alpha = 2, \beta = 0$). For this reason we consider two field patterns that differ only by an overall multiplier to be essentially the same mode.

This operator notation is reminiscent of quantum mechanics, in which we obtain an eigenvalue equation by operating on the wavefunction with the Hamiltonian. Main key properties of the eigenfunctions of the Hamiltonian are following: they have real eigenvalues, they are orthogonal, they can be obtained by a variational principle, and they may be cataloged by their symmetry properties.

In the chapters devoted to semiconductor device structures and plasmonic crystals we will deal with this eigenvalue problem extensively. Special features and techniques appropriate for each type of materials will be discussed in corresponding sections of this work.

2.2.3 Electromagnetism and quantum mechanics

There are a lot of similarities in the formalism and the equations of quantum mechanics, solid state physics and electrodynamics. Through these similarities we can relate different phenomena that take place for different quasi-particles such as photons/plasmons in photonic/plasmonic crystals to familiar phenomena for electrons in solid state physics, where on the atomic scale the electrons show wave properties. Since quantum mechanics of periodic potentials is the basic theory of solid-state physics (5; 6), the subject dealing with photonic/plasmonic crystalls can easily inherit some of the theorems and concepts from solid-state physics with minor modifications. Such inheritance is possible because the basis of photonic crystals is the propagation of the electromagnetic wave in periodic dielectric media. Quantum mechanics also deals with wave propagation, one difference being that the physical meaning of such waves is different from those in electrodynamics (7; 1).

We would like to present some similarities and differences between electrodynamics of dielectric media and the quantum mechanics of noninteracting electrons.

In both cases, we decompose the fields into harmonic modes that oscillate with a phase factor $e^{i\omega t}$. In quantum mechanics, the wave function is a complex scalar field. In electrodynamics, the magnetic field is a real vector field and the complex exponential is just a mathematical convenience.

In both cases, the modes of the system are determined by a Hermitian eigenvalue equation. In quantum mechanics, the frequency ω is related to the eigenvalue via $E = \hbar\omega$, which is meaningful only up to a constant V_0. In electrodynamics, the eigenvalue is proportional to the frequency squared, and there is no arbitrary additive constant.

One very important difference is that in quantum mechanics, the Hamiltonian is separable if $V(\mathbf{r})$ is separable. For example, if $V(\mathbf{r})$ is a product of functions $V_x(x)\,V_y(y)\,V_z(z)$, then the problem separates into three problems–one for each direction. In electrodynamics, such a factorization is not possible. The differential operator is $\mathbf{\Theta}$, which couples the different directions even if $\varepsilon(\mathbf{r})$ is separable. This makes analytical solutions more difficult. To address most of the electromagnetic phenomena, we will have to make use of numerical solutions.

Field	$\Psi(\mathbf{r},t) = \Psi(\mathbf{r})\,e^{i\omega t}$ Wave function in Bloch form	$\mathbf{H}(\mathbf{r},t) = \mathbf{H}(\mathbf{r})\,e^{i\omega t}$ Harmonic modes in Bloch form
Eigenvalue problem	$H\Psi = E\Psi$	$\boldsymbol{\Theta}\mathbf{H} = (\omega/c)^2\mathbf{H}$
Hermitian operator	$H = \frac{-\hbar^2\nabla^2}{2m} + V(\mathbf{r})$	$\boldsymbol{\Theta} = \nabla\times\left(\frac{1}{\varepsilon(\mathbf{r})}\nabla\times\right)$
Periodicity of the system	$V(\mathbf{r}) = V(\mathbf{r}+\mathbf{R})$ Periodic potential for all lattice vectors $\mathbf{R}$	$\varepsilon(\mathbf{r}) = \varepsilon(\mathbf{r}+\mathbf{R})$ Periodic dielectric for all lattice vectors $\mathbf{R}$
Variational theorem	$E_{var} = \frac{<\Psi\|H\|\Psi>}{<\Psi\|\Psi>}$ is minimized when Ψ is an eigenstate of H	$E_{var} = \frac{(\mathbf{H},\boldsymbol{\Theta}\mathbf{H})}{(\mathbf{H},\mathbf{H})}$ is minimized when $\mathbf{H}$ is normal mode of $\boldsymbol{\Theta}$

Table 2.1: Electromagnetism and quantum mechanics

In quantum mechanics, the lowest eigenstates typically have an amplitude concentrated in regions of low potential, while in electrodynamics the lowest modes have their electrical energy concentrated in regions of a high dielectric constant. Both of these statements are based on a variational theorem (E_{var} is minimized when Ψ is eigenstate of $\mathbf{H}$).

Finally, in quantum mechanics, there is usually a fundamental length scale that prevents from relating solutions to potentials which differ by a scale factor. Electrodynamics is free from such a length scale, so the solutions we obtain are easily scaled (2).

2.3 Diffraction limit

2.3.1 Kirchhoff diffraction theory

The spatial resolution of an optical imaging system is limited to the half of the light wavelength used in the system. This limit is usually called the Rayleigh-Abbe diffraction limit and closely related to the wave nature of light. Diffraction emerges when light interacts with geometrical objects with a size that is comparable with the light wavelength λ. Such interaction causes deviations of the light beam from the path predicted by geometrical optics. Until recently, the diffraction limit was considered fundamental and often related to Heisenberg uncertainty principle.

Any light source produce the electromagnetic field that consist of two components. One is the propagating FF component and other is decaying NF component. We need both to completely restore information about source. Inability to detect

NF component is the reason why diffraction limit is fundamental problem for conventional FF optic. The only way to break it is to restore information from NF component. That's mean NF field detection, but that is possible only with means of NF optic.

The first diffraction theory (Huygens-Fresnel theory) was based on the idea that the light disturbance at some point arises from the superposition of secondary waves that proceed from a surface situated between this point and the light source (8). This idea was developed by Kirchhoff, who derived it from first principles (1).

We consider a scalar field (a component of **E** or **B**) inside a closed volume V. Possible diffraction geometries for V are shown in Fig. 2.1. We assume that the scalar field has a harmonic time dependence

$$\psi(\mathbf{r}, t) = \psi(\mathbf{r})\, \mathrm{e}^{-i\omega t} \tag{2.23}$$

and satisfies the scalar Helmholtz wave equation

$$(\nabla^2 + k^2)\, \psi(\mathbf{r}) = 0 \tag{2.24}$$

inside V.

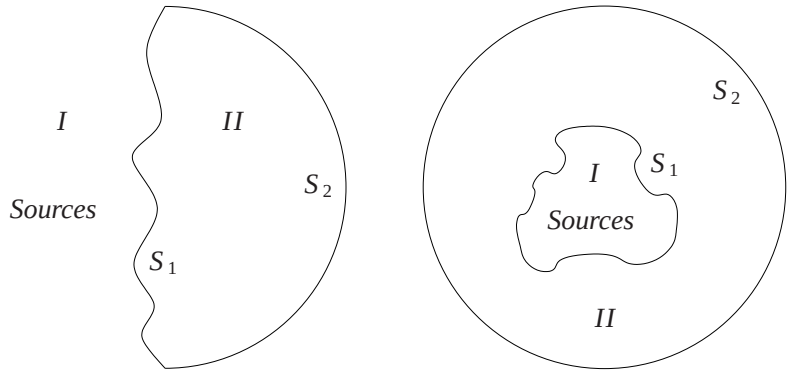

Figure 2.1: Possible diffraction geometries: S_1 is an arbitrary mathematical surface that corresponds to a screen with apertures and divides the space into two regions. Region I contains the sources of radiation. Region II is the diffraction region. S_2 is the infinitely remote surface.

By applying the Green's theorem and introducting the infinite-space Green function describing outgoing waves,

$$G(\mathbf{r}, \mathbf{r}') = \frac{\mathrm{e}^{ikR}}{4\pi R} \tag{2.25}$$

where $\mathbf{R} = \mathbf{r} - \mathbf{r}'$, Kirchhoff ends up with the formula:

$$\psi(\mathbf{r}) = -\frac{1}{4\pi} \int_{S_1} \frac{\mathrm{e}^{ikR}}{R} \mathbf{n}' \cdot \left[\nabla'\psi + ik\left(1 + \frac{i}{kR}\right)\frac{\mathbf{R}}{R}\psi \right] \mathrm{d}a' \tag{2.26}$$

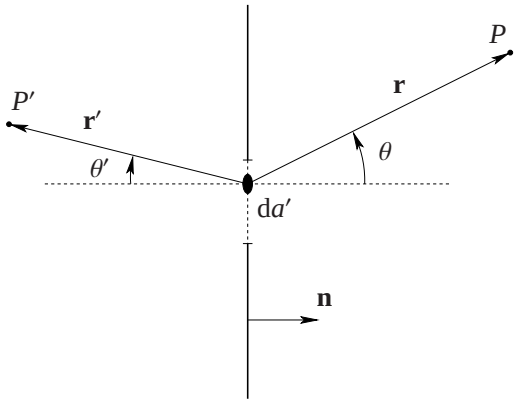

Figure 2.2: Diffraction geometry for a point source at P', a plane screen with apertures, and an observation point at P. The distances from the element of area da' in the aperture to the points P and P' are r and r', respectively. The angles θ and θ' are those between $\mathbf{r}$ and $\mathbf{n}$, and $\mathbf{r}'$ and $-\mathbf{n}$, respectively.

This is the *Kirchhoff integral* formula.

We need to know the values of ψ and $\partial\psi/\partial n$ on the surface S_1. Usually this is not the case. We can use approximate values of ψ and $\partial\psi/\partial n$ in order to calculate an approximated diffracted wave. The *Kirchhoff approximation* consists of the assumptions:

1. ψ and $\partial\psi/\partial n$ vanish everywhere on S_1 expect at the openings.
2. The values of ψ and $\partial\psi/\partial n$ in the openings are equal to the values of the incident wave in the absence of any screen or obstacles.

We consider the special case when the surface S_1 is an infinite *plane screen at* $z = 0$. Assuming a *point source* at position P' on one side of a plane screen and an observation point P on the other side, as shown in Fig. 2.2. For a spherically symmetric amplitude of the light source ($e^{ikr'}/r'$) and P, P' lying many wavelengths away from the screen we can rewrite Kirchhoff integral as:

$$\psi(P) = \frac{k}{4\pi i}\int_{apertures} \frac{e^{ikr}}{r}\frac{e^{ikr'}}{r'}(\cos\theta + \cos\theta')da' \tag{2.27}$$

2.3.2 Resolution limit of the conventional far-field optical system

We can use (2.27) to calculate the resolution limit of a conventional optical system. We consider the aperture in a screen with the radius a that is much bigger than the wavelength. Assuming that P' and P are situated at infinity we integrate (2.27) over the surface of the aperture. We can fullfill the infinity condition for P' and P with help of two lenses in the arrangement shown in Fig. 2.3. The point source P' is located at f_1, the principal focus of the length L_1, and the observation plane is the second focal plane, f_2, of L_2.

Taking this into account (2.27) becomes

$$\psi(P) \propto \int_{apertures} e^{ik\left(\frac{x}{f_2}\xi + \frac{y}{f_2}\eta\right)} d\xi d\eta \tag{2.28}$$

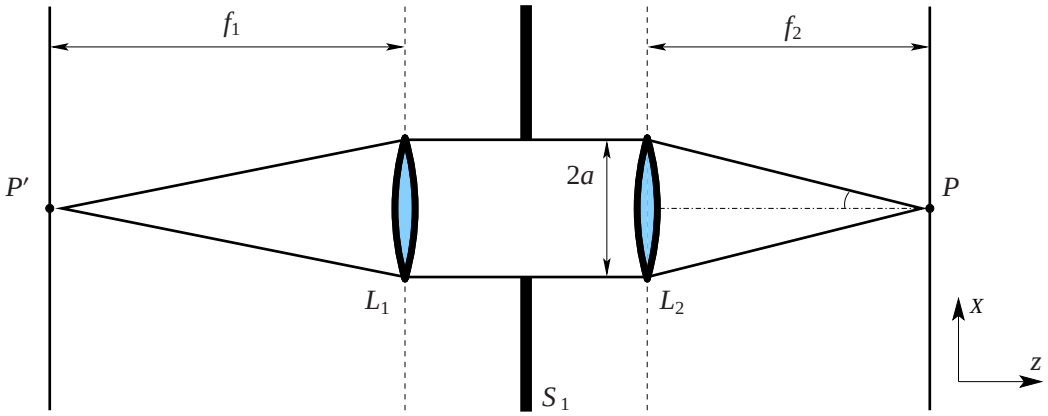

Figure 2.3: Geometry used to calculate the lateral spatial resolution of a conventional far-field microscope.

where ξ and η represent the coordinates on the screen, and x and y the coordinates in the observation plane.

The physically measurable quantity is the intensity of the electromagnetic field I. It is proportional to the squared electric field, E: $I \propto \langle E^2 \rangle$. For the case of a circular aperture with the radius a, by using the polar coordinate system and first order Bessel function, one is then able to calculate the expected intensity on the screen (9).

$$I(\theta) = I(0)\left[\frac{2J_1(ka\sin\theta)}{ka\sin\theta}\right]^2, \tag{2.29}$$

with J_1 the the first order Bessel function.

The electrical field originating from diffraction at a circular aperture and the resulting intensity distribution are shown in Fig. 2.4 (b) and 2.4 (c), respectively. The central maximum predicted by (2.29) is known as the Airy disk. It's cross section is presented in Fig. 2.4 (a). The first minimum appears at:

$$R_0 = \frac{0.61 \cdot \lambda}{NA}, \tag{2.30}$$

with $NA = n\sin\delta$ being the numerical aperture of the microscope objective, δ is the half angle of the maximum cone of light collected by this lens (see inset in Fig. 2.4 (a)), and n is the refractive index of the medium between object and lens. We can ignore the secondary maxima that appear at larger distances due to their negligible intensities.

If two objects are at a distance smaller than R_0 their images can not be resolved because in the observation plane the images will appear superimposed. R_0 it then defined as the laterally diffraction limited spatial resolution.

It is visible from (2.30) that, at a fixed wavelength, the only way to increase the lateral spatial resolution is increasing its numerical aperture. In other words one can improve the resolution by increasing the size of aperture, a, and decreasing

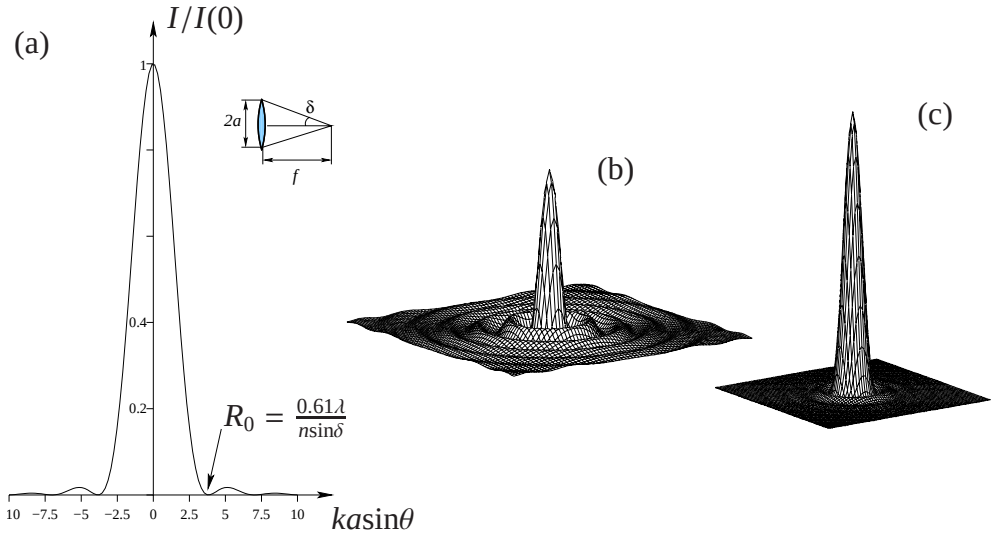

Figure 2.4: (a) The Airy pattern. (b) Electrical field created by diffraction at a circular aperture. (c) Intensity distribution resulting from diffraction at a circular aperture.

the focus distance, f, and/or by placing a high refractive medium between the objective and the sample. Commercial lenses have maximmum numerical aperture of about 0.9 for air and 1.4 for oil immersion objectives. Thus, the lateral resolution is typically limited to about one-half of the wavelength in air. For the case of GaAs/AlGaAs heterostructures the typical emission wavelength at low temperatures is around 800 nm, so that for these systems the theoretical lower limit of the spatial resolution is about 400 nm. This is the theoretical limit, but in real microscopes, spherical and chromatic aberrations should be considered, which further reduce the effective resolution.

2.3.3 Problems in breaking the diffraction limit

The resolution reached with conventional optical experiments is generally not sufficient to resolve subwavelength particularities. There are several approaches developed to improve the situation.

Folowing the discussion from the previous subsection on light diffraction further improvement may be obtained by solid immersion lenses (10) with refractive indexes of up to 3.2 (11). In any case the best resolution with has been reached so far in this way is about $\lambda/3$ (12).

Methods based on the injection of electrons such as cathodoluminescence spectroscopy (13; 14; 15) or scanning tunneling luminescence (16) have been used for monitoring the local properties of nanostructures on sub-μm lengths scale. However, in such measurements the initial energy distribution of carriers, their relaxation towards quasi-equilibrium and the transport of carriers inside the nanostructure are difficult to control and complicate the data interpretation.

A different way to perform single nanostructure spectroscopy is to reduce the optically active part of the sample. This can be done by using electron-beam lithography and metal liftoff to open a series of small holes (mesas) in a 100-nm-thick Al film which is deposited on the surface (17; 18; 19). The holes, ranging on diameter from several μm to 100 nm, are spaced sufficiently far apart to allow optical probing of a single hole with a conventional photoluminescence (PL) setup. Similarly the optically active region can be reduced in the 100 nm range by removing, with electron-beam lithography and wet chemical etching, the nanostructures except in small mesa structures, so that the sample consists of a series of small fields, each with a reduced number of nanostructures (20; 21; 22). Also in this case, the mesas, spaced sufficiently apart, are characterized by lateral dimensions between several μm and 100 nm, so that single nanostructure spectroscopy can be performed with standard PL techniques. A disadvantage of this approach is that it does not allow investigations of the spatial variation of optical properties of the sample. Therefore, it is not possible to get information about the spatial distribution of exciton wavefunctions and their localization properties. A second critical point is that it is not yet clear if metal nano-apertures on a semiconductor nanostructure are indeed of non-invasive nature. Heltzer et al. show that the semiconductor-insulator-metal contact leads to an electrostatic potential in the semiconductor that accumulates holes at the edge and electrons in the center (23). This inhomogeneous carrier distribution leads to unusual linear and non linear features in the optical spectra.

To overcome the diffraction limit of conventional optical microscopy and to push the spatial resolution of optical microscopy into the 10-100 nm regime one has to work in the NF regime. In this way one can take advantage of all the properties of optical microscopy, e.g., imaging of wide sample areas, resonant excitation and, thus, generation of carriers under well-defined conditions without being limited in determining the spatial accuracy by light diffraction.

2.4 Theoretical concepts of the near-field scanning optical microscopy

2.4.1 Hertzian dipole

We consider a system of charges and currents that vary in time. We apply Fourier analysis to the time dependence of the system, extract the individual Fourier components and treat them separately (1). In this way we can separate a complicated task into a set of simple charges and currents that vary sinusoidally in time:

$$\rho(\mathbf{r}, t) = \rho(\mathbf{r})\, \mathrm{e}^{-i\omega t} \tag{2.31}$$

$$\mathbf{J}(\mathbf{r}, t) = \mathbf{J}(\mathbf{r})\, \mathrm{e}^{-i\omega t} \tag{2.32}$$

Such a dependence is characteristic for a Hertzian dipole, i.e. an electrical dipole where the negative charge is fixed at $\mathbf{r} = 0$ and the positive charge oscillates with a linear harmonic motion around the origin.

We can write the solution of the Maxwell equation for a vector potential $\mathbf{A}(\mathbf{r}, t)$ in Lorentz gauge as follows:

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int \mathrm{d}^3 r' \int \mathrm{d}t' \, \frac{\mathbf{J}(\mathbf{r}', t)}{|\mathbf{r} - \mathbf{r}'|} \, \delta\left(t' + \frac{|\mathbf{r} - \mathbf{r}'|}{c} - t\right) \tag{2.33}$$

Here, $\mathbf{r}$ represents the position for which we want to calculate the electromagnetic field, and $\mathbf{r}'$ the position of the dipole.

With a sinusoidal time dependence (2.32) the vector potential becomes

$$\mathbf{A}(\mathbf{r}, t) = \frac{\mu_0}{4\pi} \int \mathbf{J}(\mathbf{r}') \, \frac{\mathrm{e}^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r} - \mathbf{r}'|} \, \mathrm{d}^3 r' \tag{2.34}$$

where k is the wavevector and $|k| = \omega/c$ the wave number. The magnetic field is given by

$$\mathbf{H} = \frac{1}{\mu_0} \nabla \times \mathbf{A} \tag{2.35}$$

while, the electric field is

$$\mathbf{E} = \frac{iZ_0}{k} \nabla \times \mathbf{H} \tag{2.36}$$

where $Z_0 = \sqrt{\mu_0/\varepsilon_0}$ is the impedance of the free space.

Assuming that the size d of the oscillator is small compared to the wavelength $d \ll \lambda = 2\pi c/\omega$, we can distinguish three spatial regions of interest:

The near-field zone:	$d \ll r \ll \lambda$
The intermediate zone:	$d \ll r \sim \lambda$
The far-field zone:	$d \ll \lambda \ll r$

We first concentrate on the far-field zone ($\lambda \ll r$ or $kr \gg 1$). The exponential in (2.34) oscillates rapidly and determines the behavior of the vector potential. Here it is sufficient to approximate

$$|\mathbf{r} - \mathbf{r}'| \simeq r - \mathbf{n} \cdot \mathbf{r}'. \tag{2.37}$$

Actually this is valid for all $\mathbf{r} \gg d$, independent of the value of kr. So, even for the near-field zone, it is the adequate approximation. Using this approximation we can transform the vector potential into

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \, \frac{\mathrm{e}^{ikr}}{r} \int \mathbf{J}(\mathbf{r}') \, \mathrm{e}^{-ik\mathbf{n}\cdot\mathbf{r}'} \, \mathrm{d}^3 r' \tag{2.38}$$

Applying integration by parts and using the continuity equation $i\omega\rho = \nabla \cdot \mathbf{J}$ we can rewrite the integral as follows:

$$\int \mathbf{J} \, \mathrm{d}^3 r' = -\int \mathbf{r}' \, (\nabla \cdot \mathbf{J}) \, \mathrm{d}^3 r' = -i\omega \int \mathbf{r}' \rho(\mathbf{r}') \, \mathrm{d}^3 r' \tag{2.39}$$

Thus the vector potential is

$$\mathbf{A}(\mathbf{r}) = \frac{-i\mu_0\omega}{4\pi}\,\mathbf{p}\,\frac{\mathrm{e}^{ikr}}{r} \tag{2.40}$$

where

$$\mathbf{p} = \int \mathbf{r}'\rho(\mathbf{r}')\,\mathrm{d}^3r' \tag{2.41}$$

is the *electric dipole moment.*

Knowing the vector potential we can find from (2.35) and (2.36):

$$\mathbf{H} = \frac{ck^2}{4\pi}\,(\mathbf{n}\times\mathbf{p})\,\frac{\mathrm{e}^{ikr}}{r}\left(1-\frac{1}{ikr}\right) \tag{2.42}$$

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}\left\{k^2(\mathbf{n}\times\mathbf{p})\times\mathbf{n}\frac{\mathrm{e}^{ikr}}{r} + [3\mathbf{n}(\mathbf{n}\cdot\mathbf{p})-\mathbf{p}]\left(\frac{1}{r^3}-\frac{ik}{r^2}\right)\mathrm{e}^{ikr}\right\} \tag{2.43}$$

In the far-field zone the fields take the form:

$$\mathbf{H} = \frac{ck^2}{4\pi}\,(\mathbf{n}\times\mathbf{p})\,\frac{\mathrm{e}^{ikr}}{r} \tag{2.44}$$

$$\mathbf{E} = Z_0\,\mathbf{H}\times\mathbf{n} \tag{2.45}$$

while in near-field we find:

$$\mathbf{H} = \frac{i\omega}{4\pi}\,(\mathbf{n}\times\mathbf{p})\,\frac{1}{r^2} \tag{2.46}$$

$$\mathbf{E} = \frac{1}{4\pi\varepsilon_0}\,[3\mathbf{n}(\mathbf{n}\cdot\mathbf{p})-\mathbf{p}]\,\frac{1}{r^3} \tag{2.47}$$

From equations (2.46) and (2.47) follows that the magnetic induction in the near-field region is by a factor kr smaller than the electric field. Therefore, in the limit $kr \ll 1$ the fields are predominantly of electric nature. Moreover, the near-field components, that depend on the detailed properties of the dipole decay very rapidly with distance. Consequently, in the picture where every radiation source can be regarded as an ensemble of Hertzian dipoles, the near-field components of the electromagnetic field are carrying all the information needed to determine the spatial distribution of the emitter. Thus, the main idea of near-field microscopy is to exploit these components by enhancing the spatial resolution through a very close placement of the object to be imaged to a sensing element. This sensing element should then be able to detect the non propagating evanescent field (2.46) (2.47) and to transform its spatial distribution information into the far-field. Such an evanescent field is always generated if light is diffracted by finite-size objects.

2.4.2 Resolution limits in near-field optics

There are various approaches to the problem of definition and determination of the resolution of an imaging system (24; 25; 26). The most widely used definition of resolution is as the smallest distance between two point-like objects that can still be resolved. For the far-field case we have to consider the classical Rayleigh-Abbe diffraction limit of resolution: $d = \lambda/2\sin(\theta)$. For the near-field configuration, the spatial resolution is no longer limited to one-half of the light wavelength but only by the probe-sample separation: $d \gg z_p$. Since the probe-surface distance means actually the distance between the scattering *center* of a probe and the surface, we can state that *the resolution of the NSOM is determined by the size of the probe and its distance to the surface.* On the other hand the evanescent wave can be characterized by a spatial period ($\lambda_{ev} = 2\pi/p$). The resolution limit of the NSOM is therefore determined by *the smallest period of the optical field participating in the probe-sample interaction.* The latter conclusion does not contradict with the former one but rather generalize it.

2.4.3 Experimental configurations

In any experimental NSOM configurations there are some essential components that should be present: a sample with subwavelength surface features illuminated by light; a subwavelength-sized probe that is scanned along the surface at a close distance; and a detection system that records the light scattered by the probe as a function of scanning coordinates. One can also use a subwavelength sized source of the incident light that is scanned along the sample surface and detect the light scattered by the sample. The way of practical experimental realization appears to be limited only by the imagination of near-field spectroscopists. There are, however, a handful of configurations that dominate most of the research (24; 25). Five of the most common modes are described in Fig. 2.5. Here we assumed that the signal transduction element is tapered fiber. In all cases shown in Fig. 2.5, the sample is lying on a flat surface just below the apex of the fiber. In Fig. 2.5(a-c), the sides of the fiber are coated with a metal film. Using such tips with small opening at the very end (27) represents a logical development of the original idea of using a small aperture in an otherwise opaque screen (28). In Fig. 2.5(d), no such coating exists. And in 2.5(e) a totally metallic tip is employed.

In illumination mode NSOM [Fig. 2.5(a)], the sample is illuminated with light from the aperture at the end of the fiber, and the signal is collected with far-field detector. In collection mode [Fig. 2.5(b)], light is collected through the aperture while illumination is provided by a far-field source. In the illumination/collection NSOM mode [Fig. 2.5(c)], the sample is illuminated with the light from aperture, and light from sample is collected through the same aperture. In the photon tunneling NSOM configuration [Fig. 2.5(d)], one uses confined illumination by a totally internally reflected wave incident onto the surface from the side of the sample (29; 30; 31). In this case the background field is tightly bound to the sample

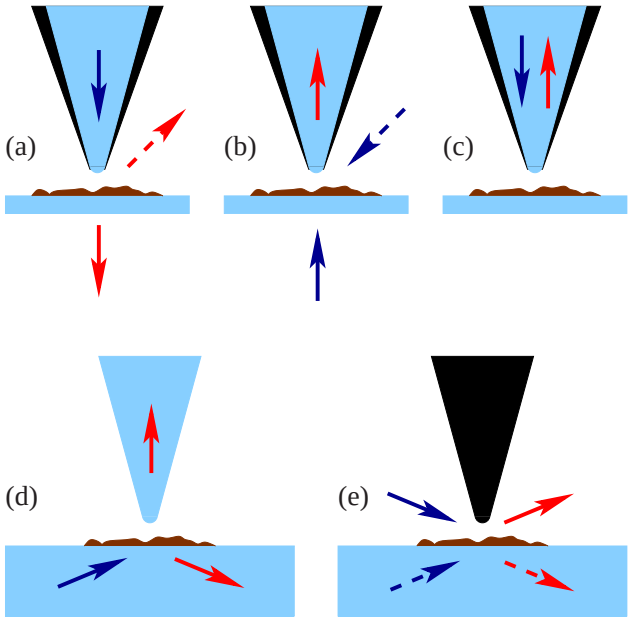

Figure 2.5: Different configurations employed in NSOM. (a) Illumination mode; (b) collection mode; (c) illumination/collection mode; (d) photon tunneling mode; (e) apertureless mode.

surface, and the sharp uncoated tip can be considered a point like probe of electric near-field intensity (32). In such a configuration, the use of an uncoated tip is preferable because of the relatively low refractive index of optical fiber decreases the risk of perturbing of the measured field.

The probes presented in Fig. 2.5(a-d) have the advantage of using a fiber to guide the light to/from the sample. Despite this there are some limiting factors. In the best case of a metal-coated aperture probe light confinement is limited by the electric field skin depth of the metal coating (e.g., ≅ 12 nm for an Al coating at the light wavelength of 633 nm). This strict limit to light confinement and the corresponding resolution limit in the aperture NSOM along with the realization that the light is also confined near sharp material boundaries has stimulated interest in *apertureless* NSOM techniques (33; 34; 35; 36; 37). This technique is highly sensitive to the polarization of the incident light. In order to realize a significant field enhancement (near the tip) the incoming field should have a large component along the axial direction (38; 39; 40). For this reason it is advantageous to use a *p*-polarized light beam (electric field in the plane of incidence) incident under an oblique angle either from the side of the tip or from the side of the sample [Fig. 2.5(e)] for illumination.

Chapter 3

Experimental setup

This chapter will be devoted to the experimental technique. Only general concepts common for all type of our experiments will be described. Details that are characteristic for the semiconductor device structures or plasmonic crystals investigations will be discussed in the corresponding chapters.

In any NSOM experiment the NF-tip is playing a key role. So we will start our survey from the different types of the tips and their evolution with maturing of the NSOM technique. Different technologies for fiber tip fabrication will be described. Advantages and disadvantages of each of them in resolution vs. transmissivity trade-off will be shown.

The fiber tip is the most sensitive part in a NSOM setup. During scanning it should be in close vicinity of the object under investigation, but any contact with the sample will destroy it. To avoid contact and to keep the tip close to the surface a feed-back mechanism is employed. Its main principles and realization possibilities will be discussed. The optical feed-back system used in our experimental setup will be described.

The description of a room temperature NSOM in one of the common configurations will be performed. Laser sources with fixed and tunable wavelengths will be described. We will discuss different types of the detection techniques that were employed in the scope of this work.

3.1 Fiber tips and resulting quality of the experiment

3.1.1 Main types of near-field tips

The near-field tip is the key element of the NSOM. As the technological development allowed producing in controllable way a subwavelength aperture at the tip end, the practical implementation of Sygne's ideas (28; 41) became possible.

Different types of NF-tips have been developed and presented in literature. The first demonstration of NSOM imaging at visible wavelengths (42) involved

the use of a cleaved quartz crystal that was anisotropically etched and then coated. The next step was the use of pulled glass pipettes (43; 44). The metall-coating process ends with forming a small aperture at the very end of such a pipette. As near-field microscopy has matured, an increasing fraction of NSOM imaging is undertaken using the coated tapered optical fiber as suggested by Betzig et al. (45) and followed by others (46; 47; 48). Another interesting method, because of its high reproducibility, is microfabrication of the near-field tips similar to those ones used in scanning-force-microscopes (49). Until now the most commonly used tips are prepared from single-mode glass fiber, both pulled (45) and etched (46).

Pulling is a simple way to produce high quality NF-tips. The fiber is locally heated with a CO_2 laser and drawn to a sharp taper in a commercial micropipette puller (e.g. Sutter Instruments, P-2000). The resulting taper has an approximately conical shape with an opening angle of typically 10° and a flat end face. The diameter of the flat end face (<30 to 300 nm) and the exact geometry of the taper can be varied by careful adjustment of the pulling parameters (50). After angeled evaporation of the sides of the taper with ~100 nm of aluminum, apertures ranging from <50 to >500 nm in diameter are reproducibly formed. With such probes, a resolution of 12 nm ($\lambda/43$) was claimed from scans of single aluminum particles on a glass substrate (45). The transmission efficiencies of pulled tips are typically on the order of 10^{-5} for an aperture diameter of 100 nm and decrease to less than 10^{-6} for diameter less than 50 nm (51), limited by the small taper angle of about 10° and the correspondingly long evanescent field region inside the taper, where the fiber diameter decreases below the wavelength of the transmitted light. The maximum power that can be launched into these probes is mainly limited by light absorption and heating of the metal coating, resulting in tip destruction for incident powers of more than a few milliwatts (52; 53; 54; 55). The polarization of the incoming light is mainly preserved within these fibers (56; 57; 58). The properties, as well as the manufacturing and coating of the pulled fiber tips, have been carefully investigated (59; 50; 47). Recent attempts to form the apertures in pulled fiber probes in a more controlled way rely on focused ion beam milling or drilling (58; 60; 61). With this method, aperture diameters as small as 20 nm are formed and an excellent behavior is obtained.

Strategies for increasing the probe transmissivity rely on increasing the taper angle in order to decrease the evanescent field region within the taper. Fiber probes with cone angles around 30° were fabricated by wet chemical etching of the fiber (62; 63; 64; 65). Transmission efficiencies of about 10^{-3} were reported for fiber giving a spatial resolution of 100 nm, an improvement by two orders of magnitude over pulled fiber probes. Saiki et al. (51) proposed a wet chemical etching in buffering condition that results in double-tapered structures. The metal cladding region with strong optical losses is then shortened by pounding the metal-coated probe on the sample. Spatial resolution of <30 nm was demonstrated. For probes giving a spatial resolution of 150 nm the typical transmission efficiency is about 5%. It was found that in general the surface roughness of the

etched probes and thus the reproducibility of the manufacturing process can be improved by performing the etching through the plastic jacket of the fiber (66).

3.1.2 Fabrication of near-field fiber tapers

The subwavelength aperture at the very end of the near-field tip determines the spatial resolution of most modern NSOMs. But the quality of the data that can be obtained depends not only on the lateral resolution but also on the transmissivity of the aperture/tip-apex system. The smallest aperture can be created with help of the pulled tips but they have a small opening angle (~10°) and a low transmissivity. This fact lead to significant experemental difficulties, because one is forced to work with extremely small signals. In order to improve the tip transmissivity one should produce large opening angles, that determine a strong decrease of the loss region inside the taper. There are a number of different methods proposed to improve this parameter (see previous subsection). All experiments on semiconductor device structures presented in this work are performed with home made tapers that are created by etching. This method ensures probes with a large aperture cone angle. And differently from etched probes the surface roughness on the taper is drastically reduced, so that the evaporated metal coating is virtually free of sides holes.

The used mechanism was first proposed by Lambelet et al. (66) and Stökle et al. (67) and is based on chemical etching of a glass fiber through its acrylate jacket, as opposed to the standard chemical etching usually done on bare fibers after removal of the jacket. The etching setup is shown in Fig. 3.1. A single mode

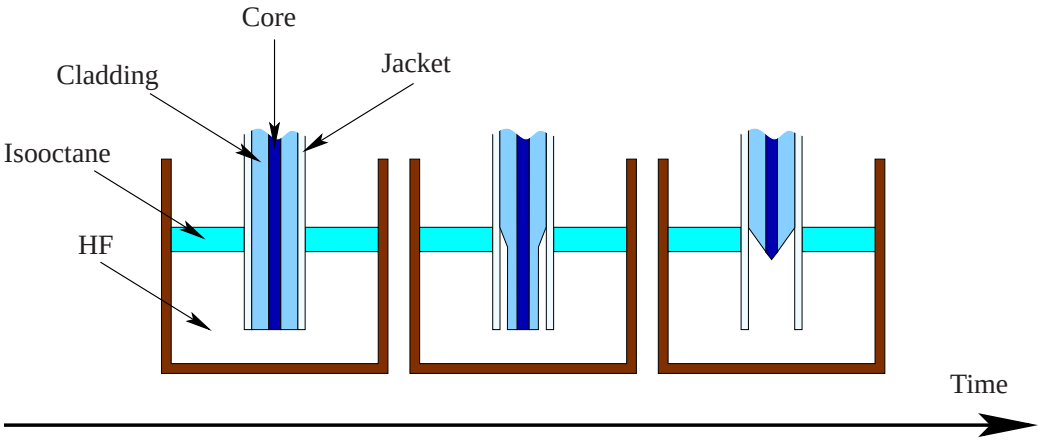

Figure 3.1: The etching setup.

optical fiber (3M, SN 4224) is dipped with its acrylate jacket into an aqueous 48% hydrofluoric acid (HF) solution. The HF solution is covered by an organic overlayer (isooctane) to protect the fiber against acid vapor. The acid does not dissolve the acrylate jacket, but rather diffuses through it to etch the quartz fiber.

90 minutes later the fiber is removed and successively rinsed with ethanol. At this point, a tip has been formed inside the acrylate jacket. To remove the jacket, an incision is made a few millimeters above the tip and the jacket softened by dichloromethane can then be pulled by seizing it in front of the tip. The jacket is thus removed without damage to the fiber, after the tip is formed. The surface of the tips etched through the jacket is very smooth and does not suffer from the surface irregularities usually encountered in chemically etched tips. This improvement of the quality of the surface is due to the different physical processes responsible for the formation of the tip. When etching the bare fiber, the meniscus height decreases as the fiber diameter is reduced by the etchant during the process. This process is irregular due to surface tension forces. With a protected fiber, the height of the meniscus is governed by the interface acid-acrylate-organic layer. As the acrylate jacket is not attacked by the acid, this interface is much more stable, and in any case does not influence much the tip formation which is mostly governed by concentration gradients and micro-convection inside the tube. Moreover, since the whole etching process takes places inside a hollow cylinder formed by the protective acrylate jacket, it is less sensitive to environmental influences such as vibrations, temperature drifts, and so on. If tips etched through the jacket are subsequently coated with an aluminum layer of about 100 nm the coating shows a good homogeneity and leakage of light along the cone of the taper can be avoided. The probes are coated in a high vacuum evaporator (Edwards high vacuum, AUTO 306, thermic vaporizer). During the aluminum evaporation the probes rotate along their axis, so that at the end of the tip a circular aperture is formed. The aperture diameter depends on the inclination of the probes in respect to the aluminum source, on the thickness of the metal layer and on the dimension of the plateau at the end of the taper.

3.2 Feedback mechanism

3.2.1 NSOM feedback basics

The resolution of the NSOM is strongly influenced by the distance between NF-probe and surface of the sample. The distance should not exceed a small fraction of the light wavelength used in the experiment. In case of the visible spectral range it should not exceed a few nanometers. It was proved that an increase in distance between NF-tip and sample surface causes a change in field distribution near the tip and causes contrast losses in the detected images (e.g. (27)). Tips that are used in the NSOM are extremely sensitive to impacts caused by mechanical forces. The distance control mechanism should keep the near-field tip close (but contact-free) to the scanned surface. The control mechanism also should be able to perform a crash-free approach process. It should observe and navigate the tip through the way to the sample surface.

One can find in the entire literature very different approaches to solve the

distance control problem. In earlier works it was proposed to control the tip-sample distance with help of the current between metalized tip and conducting sample (42). Later this method was improved by Dürig et al. (68) through the tunnel-current implementation from the principles of a tunnel-microscopy. This method is limited to conducting samples only. Other authors suggest optical methods, such as the "photon tunneling" method (30; 29), the measurement of the "background fluorescence increase" (69) or the measurement of the "near-field reflection"(70; 71).

The other way to solve the distance control problem became clear after the discovery of the shear-force damping and the development of the shear-force distance-control mechanism by Betzig et al. (72) and at the same time by Toledo-Crow et al. (73).

The authors' observations show that a laterally oscillating NF-tip, excited in perpendicular direction to the normal of the surface, behaves like a harmonic oscillator experiencing viscous damping by the water layer at the surface of the sample. Regardless to the source of the damping the motion equation of the tip can be presented as follows (74):

$$m\frac{\partial^2 x(t)}{\partial t^2} + g\frac{\partial x(t)}{\partial t} + k_F x(t) = F_a \sin(\omega t) \tag{3.1}$$

where m is the effective mass of the tip, g is the damping coefficient and k_F is the field constant of the tip. The known solution of the 3.1 is $x(t) = A(\omega, Q)\sin(\omega t + \varphi)$, where the amplitude is:

$$A(\omega, Q) = \frac{F_a/k_F}{\sqrt{\left(1 - \frac{\omega^2}{\omega_0^2}\right)^2 + \left(\frac{\omega}{Q\omega_0}\right)^2}} \tag{3.2}$$

with ω_0 as the non-damped (eigen-) resonance frequency of the tip and $Q = \omega_0/\Delta\omega$ as the quality factor of the tip oscillation. The quality factor is dependent of tip design and, because, the damping wile the tip approach to the sample surface, it is a function of the distance h. In Fig. 3.2 we show the dependence of the tip amplitude on excitation frequency for different quality factors near the resonance frequency (Eq. 3.2). The oscillation amplitude at the resonance frequency ω_0 is a linear function of the quality factor Q.

The damping force given by the damping coefficient g lies in a region of few nanonewtons (75). Until now there is no common base in a controversial discussion about the physical causes of this force. The typical acting range of the damping force that was experimentally measured shows a strong dependence on the sample material and lies in the region of few nanometers (76). The possible explanations are van-der-Waals interaction, capillary force (73; 77), surface charge effects (76) between near-field tip and sample surface. Other authors are discussing purely geometrical-mechanical interaction between NF-tip and sample surface (tip "tapping" on the sample) (78). There are different sources that describe operational modes of the distance controlling system aimed to operate at

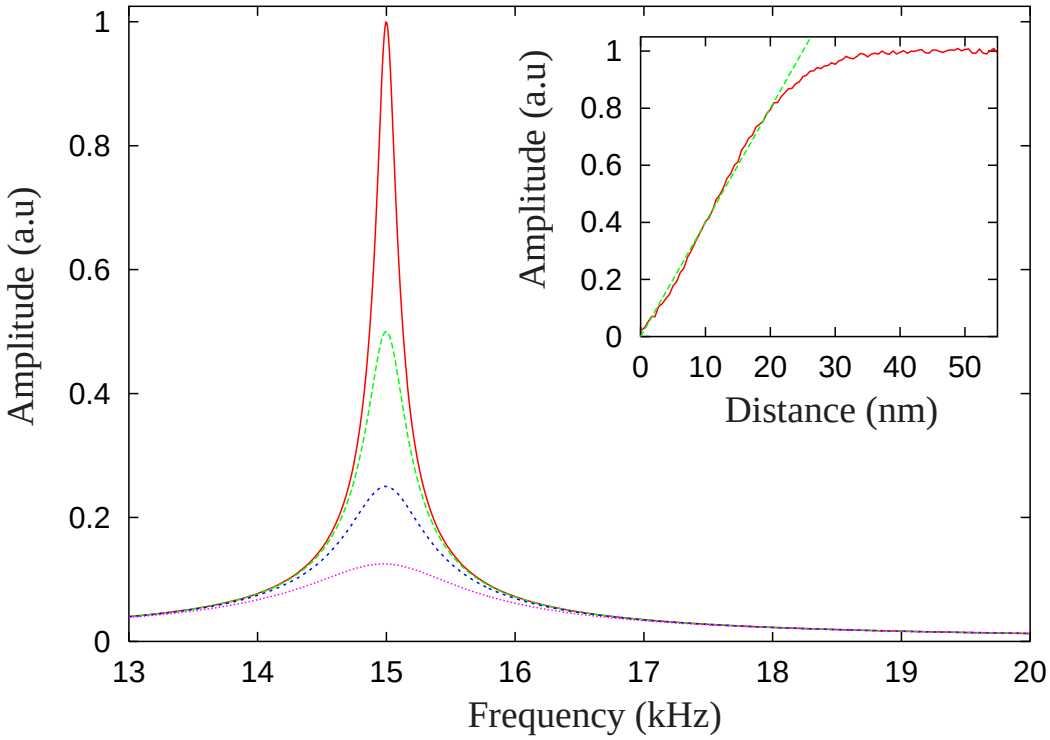

Figure 3.2: Oscillation amplitude of the NF-tip as function of excitation frequency in accordance with equation 3.2. The resonance frequency of the fiber probe is 15 kHz. The quality of the resonance is reduced from $Q = 100$ (solid line), through $Q = 50, 25$ to (dotted line). The dependence of the oscillation amplitude on the resonance frequency is a linear function of the quality factor Q. *Inset:* Approach curve of the fiber tip. The amplitude shows an almost linear behavior for $h < 22$ nm (dashed line).

different conditions. The exact design can slightly differ in dependence on the operational conditions, e.g. in atmosphere, in aquatic environment (79), in liquid helium (80) or in vacuum (76).

The shear force distance control mechanism makes it possible to measure the distance between sample surface and near-field tip and to generate from it topographycal information. In this way one can succeed in collecting an optical signal and simultaneously information on topography that is independent on optics.

The almost linear dependence of amplitude A on distance h, cf. Fig. 3.2 inset, can be used as an ideal control variable for control circle that governs the distance between tip and sample. The maximum amplitude of the vibrating tip does not exceed a few nanometers (81; 82). This fact puts a very strong demand on the accuracy of the amplitude measurement. Different techniques have been studied for sensing the dither amplitude such as optical (56), capacitance, impedance, piezoelectric (82) detection. In this work we use an optical system. It provides excellent accuracy and integration into the room-temperature NSOM is possible.

3.2.2 Optical feedback

In all experiments presented in this work we use an optical feedback system. It's implementation scheme is presented in Fig. 3.3. A 670 nm emitting laser diode

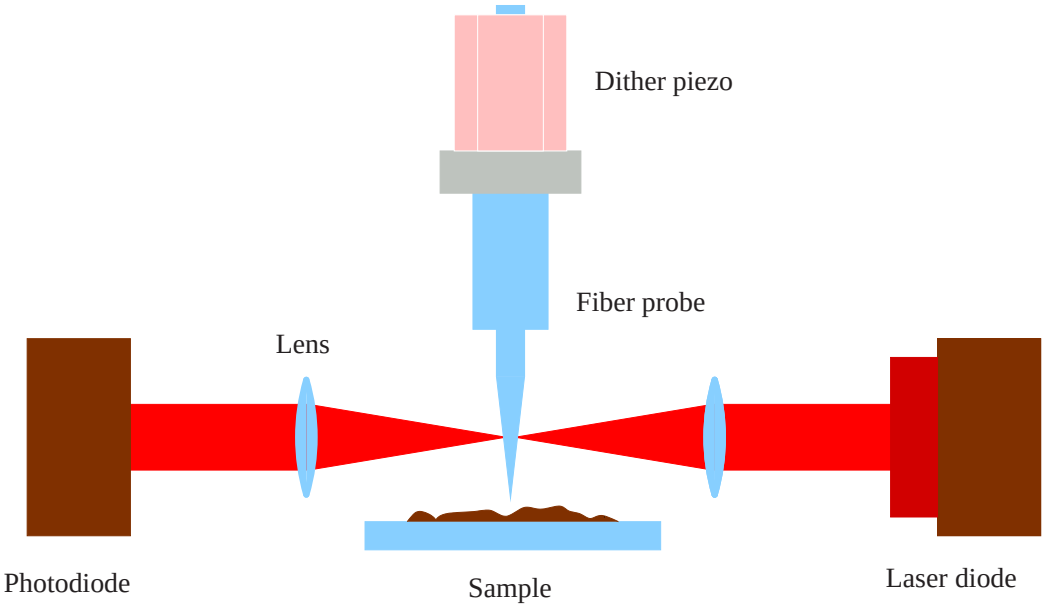

Figure 3.3: Optical part of feedback mechanism.

serves as light source whereas detection is done by a 4-segmented photodiode. The laser beam that is shaped by a $f = 50$ mm lens and a 1:3 telescope is focused on the fiber tip. The tip is vibrating in perpendicular direction to the laser light. This vibration is excited and controlled through the dither piezo. The excitation

frequency used is close to the resonance frequency of the NF-tip. The resulting deflection picture formed behind the tip is complementary to the one originated from vertical slit (in accordance with Babinet's principle (1)). With the help of a second $f = 50$ nm lens the diffraction image is projected on the 4-segmented photodiode which serves as a detector. The measured voltage difference between the segments of the photodiode is proportional to the amplitude of the tip vibration $x(t)$. With the help of a lock-in amplifier the voltage is phase-sensitively amplified. The output voltage of the lock-in amplifier is proportional to product of the tip oscillation amplitude $A(\omega)$ and the cosine of the phase shift $\Delta\varphi$ between the driving oscillation and tip oscillation ($A(\omega)\cos(\Delta\varphi)$). Then the output voltage is used as a control voltage for a PID-controller which by employing piezo-actuator is controlling the distance between the near-field tip and the samples surface. When the system is working properly it shows an extremely low noise level (tip-sample distance fluctuations are smaller than one nanometer) and a high long-time stability (more than 24 hours).

In the inset of Fig. 3.2 the approach curve (oscillation amplitude in arbitrary units as the function of distance) of the tip under a glass surface is shown. One can note that the shear-force effectively acts on distances less than 35 nm. It is clear that for $h < 35$ nm the decreasing of the tip-sample separation causes a decrease of the oscillation amplitude. After a transition region where one can see a nonlinear behavior of the approach curve one enters the region ($h < 23$ nm) where the oscillating amplitude falls linearly with distance. As it follows from oscillator model (Eq. 3.2) this region can be modelled with linear function that depends on h through a proportionality coefficient $\Delta A/\Delta h = 0.9$.

3.3 NSOM spectrometer

3.3.1 Room temperature near-field microscope

In order to perform the NF-experiments described in this work we use a room temperature NSOM. It is originally based on commercial Aurora NSOM but with a number of modifications designed to increase the performance of the device.

The central part of any NSOM is a scanning head that is responsible for the scanning process. It consists of a tip holder containing a dither piezo and a scanning stage that serves as a mount for the samples. The scanning process consists of a relative motion between the NF-tip and sample and a tip-sample distance control. In our case the tip is fixed so that it can not perform any lateral movement. The scanning head allows only a rough tip positioning in vertical direction. The unit responsible for relative tip-sample movement is the piezo scanning stage. This stage, in connection with the controll electronics, allows nanometer-precise lateral movement. The same unit sustains tip-sample distance control by employing a feed-back loop. A schematic diagram of the NSOM that operates in excitation mode is presented in Fig. 3.4.

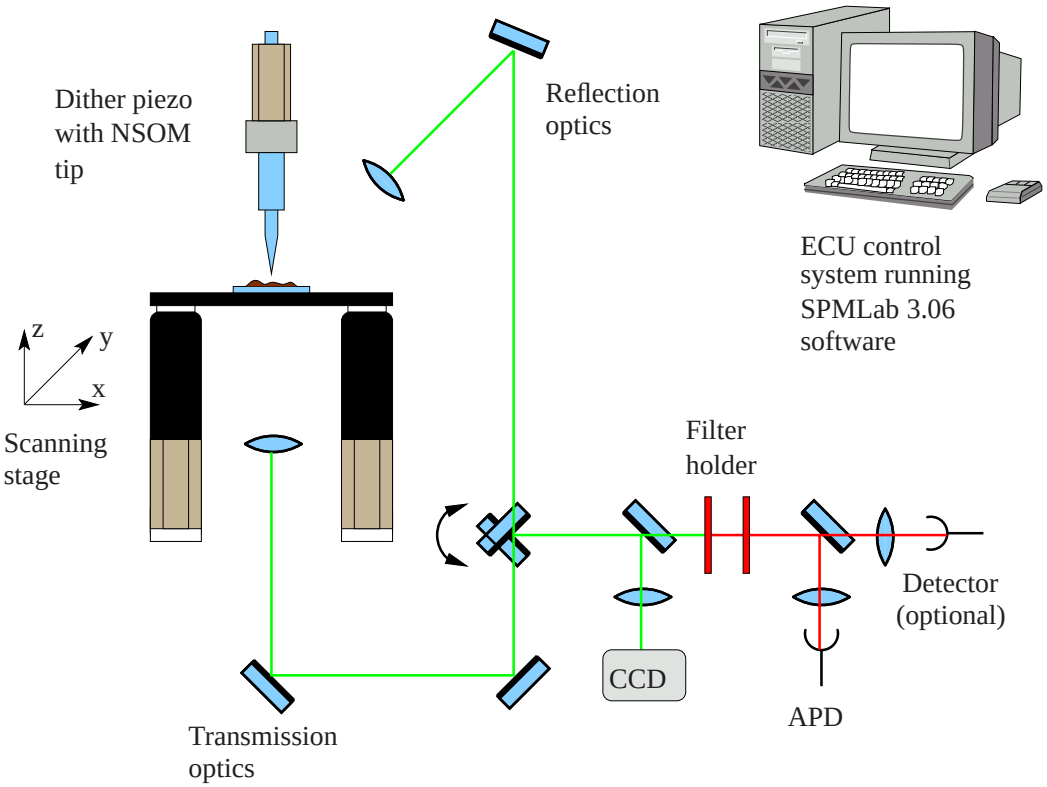

Figure 3.4: Schematic diagram of the near-field microscope.

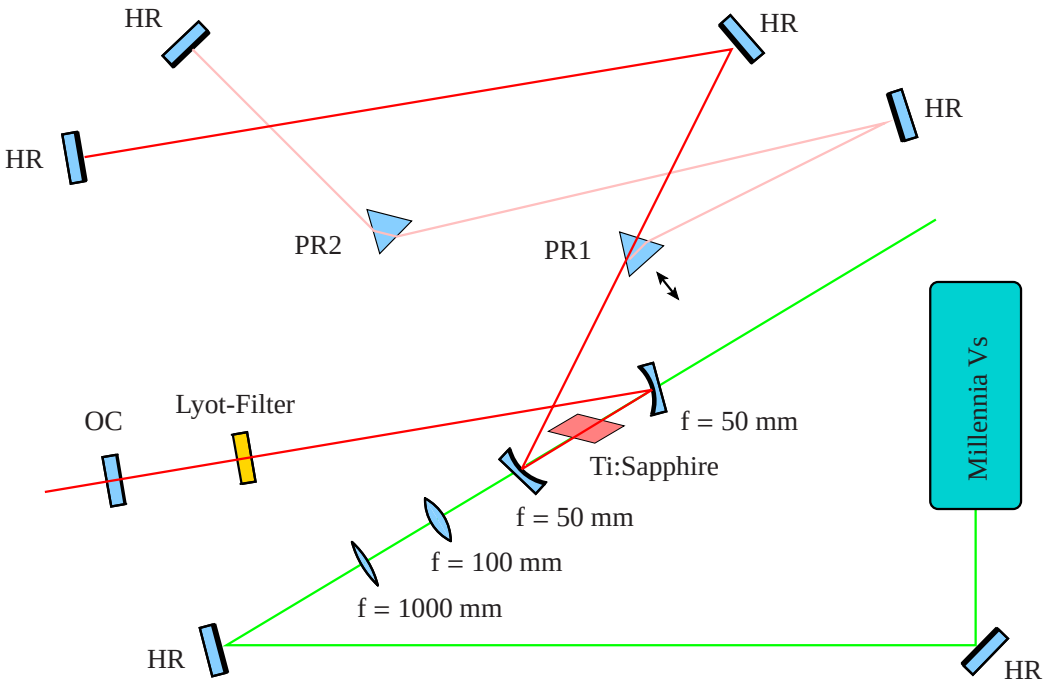

Figure 3.5: Schematic representation of the Ti:sapphire laser setup. The green, red, and pink lines show the beam path of the pump laser, the continuous wave resonator, and the short pulse resonator, respectively.

3.3.2 Excitation sources

Different types of laser sources are employed during the experiments. Commercial Helium-Neon (633 nm, 1 mW) or Cadmium-Helium (422 nm, 40 mW) lasers provide the excitation light. These lasers are used for the investigation of semiconductor device structures. In the experiments with plasmonic crystals we use a home-made tunable Ti:sapphire laser. It's schematic diagram is shown in Fig. 3.5. The Ti:sapphire crystal is pumped by a Millennia Vs (Spectra-Physics) solid-state laser (532 nm, >5 W). The laser active medium, a titanium doped sapphire crystal, with a wide fluorescence spectrum from 600 to 1100 nm offers an excellent excitation source for resonance investigation of semiconductor device structures and plasmonic crystals. The design is similar to that of Krausz et al. (83) and Ashworth et al. (84). The laser configuration consists of a 4-mirror resonator with one folded section. The optical resonator encloses two curved dichroic mirrors, which are transparent for the pump light and highly reflective for the laser radiation. Two flat mirrors close the linear cavity, one is acting as high reflector and another as output coupler. The two arms extending from the folded section are asymmetric. The longer one encloses the dispersion compensation prisms. Depending on the position of the prism PR1, two possible configurations can be chosen. If the prism is retracted, the laser runs in a continuous wave regime. If the prism inserted in the resonator, the beam follows the path defined by the Ti:sapphire crystal. To avoid

the accumulation of heat in the active medium crystal, the supporting mount is cooled with water. In continuous wave regime, when pumped by 5 W and using a 10% output coupler ($T = 10\%$) the laser produces around 400 mW of output power. The wavelength tuning is achieved by a three-plate birefringent filter inside the laser cavity. Depending on which mirror sets is mounted the radiation wavelength can be tuned in the ranges 720-810 nm, 780-880 nm and 840-920 nm, respectively.

3.3.3 Detection techniques

The near-field PL signal collected either through the tip or by a far-field microscope is dispersed by a grating and then its relative intensity is measured with appropriate detectors. PL spectra presented in this work have been recorded with a back illuminated deep depleted charge coupled device (CCD) (Princeton Instruments, LN/CCD-100EHRB detector head), in conjunction with a single grating $f = 50$ cm monochromator (Princeton Instruments, ISP-500). The CCD allows to record in a single shot at every tip position a PL spectrum with very high sensitivity. Its quantum efficiency approaches 90% in the range of maximum sensitivity. The CCD detector head is cooled with liquid nitrogen in order to reduce the background noise due to the dark current. The dark current of the CCD is 0.01 electrons/pixel-second at an operating temperature of 170 K. The spectral slit width obtained with this 1340X100 pixel CCD in conjunction with the single grating (1200 lines/mm) monochromator is 100 μeV. Therefore the combination of the near field microscope with this detection scheme allows to investigate the local optical properties of semiconductor device structures with a combined spatial and spectral resolution of 100 nm and 100 μeV, respectively.

It is not allways necessary to have such a high spectral resolution. Optionally the experiment is significantly simplified by using a fixed-wavelength detection scheme. Light collected with the NF-tip is passed through band-pass filters (spectral width ~20 nm) in order to separate the useful signal from the background. Then the PL is detected with help of a silicon avalanche photodiode.

For the device investigations performed by using the photocurrent technique the photosignal is excited through the NF-tip inside of the diode. Now the diode serves itself as the detector and the signal is detected by using the lock-in technique.

Chapter 4

Spectroscopy of semiconductor device structures

This chapter will be devoted to NF investigations of semiconductor device structures. Before presenting experimental results we will make short outlook into the growth technology of the device structures. Than we will consider emission and absorption processes in semiconductors. Basics elements of a semiconductor diode laser will be discussed. Optical gain and the threshold condition in quantum well structures will be considered.

Eigenmodes of the semiconductor laser waveguide will be considered. The method of the fake resonator for waveguide mode calculation will be introduced.

We will demonstrate the potential of near-field photocurrent spectroscopy for direct imaging of mode profiles of submicron-sized waveguides in optoelectronic devices. The technique combines the submicron spatial resolution of near-field optics with tunable laser excitation, allowing for selective investigation of the waveguide properties of the device structure. We will present experiments on InGaAs/AlGaAs high-power laser diodes with different waveguide designs that provide direct visualization of the effect of the waveguide design on (i) the number of guided modes and (ii) the spatial profile of both fundamental and higher-order modes. We will show that this technique provides a sensitive tool for nondestructive analysis of waveguide properties in optoelectronic devices.

We will present experiments performed on the monolithically stacked diode lasers, so-called Nanostack®[1] devices that include two nominally identical waveguide segments separated by a specially designed tunnel junction. We will show that near-field optical microscopy provides straightforward and separate access to the properties of both optically active segments. We will show that reduced laser emission from the laser segment that is situated closer to the substrate is not caused by thermal effects but most likely due to a larger trap concentration within or in the vicinity of the quantum-wells of this laser segment. Furthermore, we will show that in the unbiased devices the potential gradient in this segment is signif-

[1]Nanostack® is a registered trademark by OSRAM Opto Semiconductors GmbH. All rights reserved.

icantly larger than in the one close to the heat sink. In addition it will be shown that the coupling between both waveguides is marginal.

We will present a detailed analysis of the carrier diffusion near diode laser facets. A near-field microscopic technique was used to probe photoluminescence from the edge-area of a quantum well (QW). This corresponds to the solution at the facet of a quantum well laser. Near the edge, surface recombination (SR) gives rise to a gradual variation of the PL signal on a micrometer length scale. The overall shape in this transition region strongly depends on the excitation intensity. We will show how it is possible to obtain the surface recombination velocity (SRV) and the diffusion length independently. We will deduce from the experiment that the SRV is decreasing with increasing intensity due to the saturation of trap-like defect states.

4.1 Laser diodes and their epitaxial structures

4.1.1 Fabrication methods for semiconductor device structures

Since the first appearance of the semiconductor diode lasers in the earlier sixties (85) there is continuous progress in this field (86). Now the threshold-current density is more than three order of magnitude less than in early lasers. All this significant development is tightly connected to advancements in crystal-growth technologies.

The first semiconductor laser was created by Robert N. Hall at General Electric Research Laboratory in 1962 (87). It was very simple diffused laser that can lase only at helium temperatures. The first technique employed to grow more advanced diode lasers was Liquid-Phase Epitaxy (LPE) (88). This method is based on the precipitation of the semiconductor material from the melt or from the solution onto a substrate. It is a cheap method that is still widely used for mass-production of the Light-Emitting-Diodes (LED). But it has been replaced for growth of the active region of diode lasers, because of melt-back effects that hinder to fulfill needed homogeneity and thickness requirements.

Double Heterostructure (DH) diodes were succeeded from developing Chloride- and Hydride-Vapor-Phase Epitaxy (ClVPE and HVPE) (89). In this method group-III metals are supplied via reaction of a melt with chlorine in the source area with the reverse reaction occurring of the substrate. In the case of ClVPE, thrichloride group-III sources ($AsCl_3$, PCl_3) are passed over the metal melt generating the chlorine necessary for transport and at the same time providing the group-V material. In HVPE, As and P act as group-V sources and HCl is injected to transport the metal. VPE is no longer in use in diode laser production due to impossibility to make fast switching from one composition to another (because metal transport mechanism), and corrosive nature of $AlCl_3$ (which etches quartz) that make AlGaAs growth very difficult.

Two method that now are widely used for laser diodes growth are Molecular

Beam Epitaxy (MBE) and Metalorganic Chemical Vapor Deposition (MOCVD).

Most of the samples investigated in the scope of this thesis was grown with the help of MOCVD. In this technique (90; 91; 92; 93) metalorganic precursors, usually the trimethyl compounds of Al, Ga and In like $Ga(CH_3)_3$ (TMGa, also named TMG) react with sources of As and P, usually the hydrides AsH_3 and PH_3, on the heated substrate to form the epitaxial layer. This process is schematically shown in the top of the Fig. 4.1. The reaction is carried out in a gas flow with hydrogen being the usual carrier gas at atmospheric of reduced pressures. Parts of the reaction already take place in the gas phase. The final step, i.e. release of the constituent elements and incorporation into the lattice, happen on the semiconductor surface.

A typical MOCVD growth system is shown on the bottom of Fig. 4.1. It consists of three main parts – a gas-mixing system, the reactor, and the pump and exhaust handling system. In the gas-supply part the metalorganic precursors are kept in stainless steel cylinders at a precisely controlled temperature and fixed pressure. Flowing hydrohen through the liquid TMG(a), TEG(a), TMA(l), DMZn or solid TMI(n) precursors in the bottle results in a controlled transport of the saturated vapor into the reactor. The fluxes of hydrogen and the other gaseous species like AsH_3, PH_3 or SiH_4 for *n*-doping are metered by mass-flow controllers. These devices measure the thermal conductivity of the flowing gas. The pressure is controlled by capacitance of piezoelectric pressure transducers. Additional purge flows introduced downstream of the dosing devices increase the flow velocity of the reactants and thus reduce transient effects that occur when fluxes are changed (94). The reactants are switching manifolds (95). These valves are constructed for rapid switching operation and minimized dead volumes that could lead to memory effects.

4.1.2 Emission and absorption in semiconductors

In order to understand the fundamental aspects of diode lasers we consider some basic physics of the emission and absorption processes that take place in semiconductors. The important particularity of semiconductors is that that they have no sharp energy levels. The levels are broadened and form energy bands due to the overlap of the atomic orbitals. In an undoped semiconductor with no external excitation at $T = 0$ K, the lowest empty energy band is called the conduction band, and the energy band below the conduction band, called the valence band, is completely filled with electrons. These two bands are separated by an energy gap with E_g.

There are two types of carriers that contribute to electronic conduction:

- electrons in the conduction band;
- holes (missing electrons) in the valence band.

A free electron has a kinetic energy of $E = p^2/(2m_0)$ where m_0 is the free-electron mass and p the electron momentum. According to quantum-mechanics, the mo-

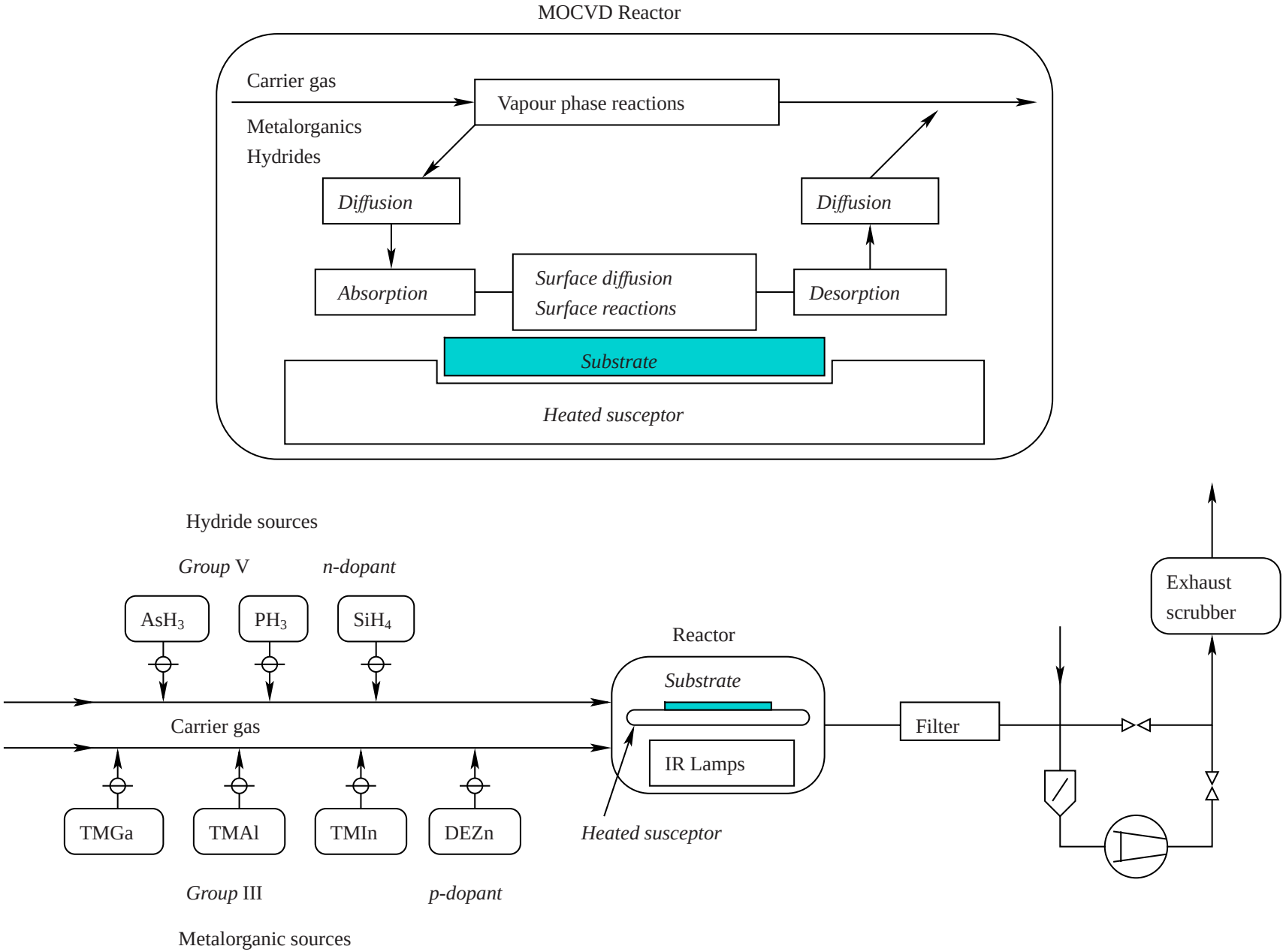

Figure 4.1: *Top:* reactions that take place during MOCVD; *bottom:* typical MOCVD system.

mentum of the particle $p = \hbar k$ is proportional to the wavenumber $k = 2\pi/\lambda$ with the reduced Plank constant $\hbar = h/(2\pi)$ and the wavelength λ. Thus the dispersion relation for a free electron is $E(k) = (\hbar^2 k^2)/(2m_0)$. In semiconductors, the electron energies in the conduction band $E_c(k)$ and in the valence band $E_v(k)$ behave similarly for small k.

$$E_c(k) = E_g + \frac{\hbar^2 k^2}{2m_e}, \qquad E_v(k) = -\frac{\hbar^2 k^2}{2m_h}. \tag{4.1}$$

This behavior, which is called the free carrier approximation, is illustrated in Fig. 4.2. The interaction of the carriers with the solid-state lattice is taken into

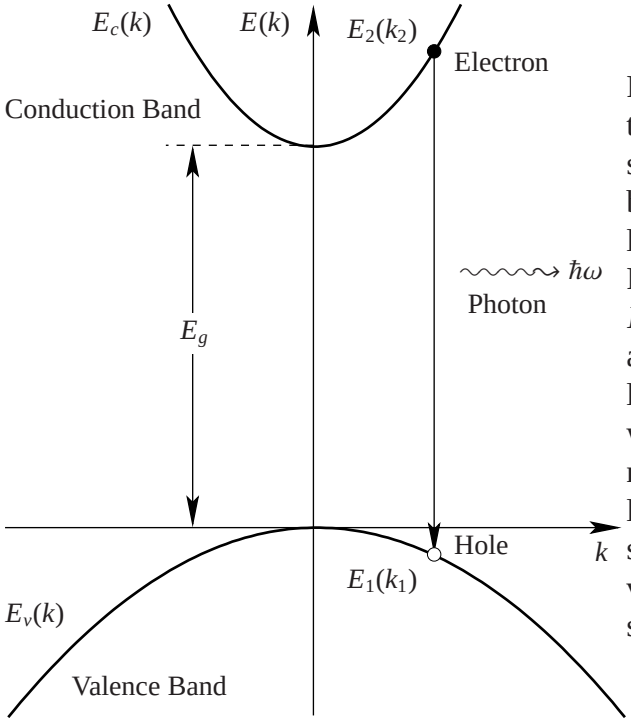

Figure 4.2: Parabolic band structure $E(k)$ for electrons in a direct semiconductor. The conduction band is separated from the valence band by an energy gap E_g. Recombination of an electron at $E_2(k_2)$ in the conduction band and a hole at $E_1(k_1)$ in the valence band generates a photon with energy $\hbar\omega$. Since the momentum of the photon $\hbar\omega$ is negligible, radiative electronic transitions between conduction and valence band only occur at the same wavenumber k.

account by the introduction of effective masses for the electrons m_e and for the holes m_h. Since the $E(k)$ dependence on the valence band is a negative parabolic curve, holes can be regarded as quasi-particles with a positive charge.

Radiative band-to-band transitions are the generation and recombination of electron-hole pairs associated with absorption or emission of photons. For these transitions, conservation of energy E and momentum $\hbar k$ must be fulfilled. Due to the high value of the speed of light c, the momentum of the photon $\hbar k = \hbar\omega/c$ for photon energies $\hbar\omega$ in the 0.5-2.5 eV range can be neglected in comparison to the momentum of the electronic carriers. Thus, a radiative transition between an electron in the conduction band with energy $E_2(k_2)$ and a hole in the valence band with energy $E_1(k_1)$ under emission or absorption of a photon can only occur at the same k value.

$$\hbar\omega = E_2 - E_1, \qquad k_2 = k_1. \tag{4.2}$$

Such a transition is shown in Fig. 4.2 by a vertical arrow with the length of the photon energy $\hbar\omega$ directed downwards for recombination (upwards for generation) of an electron-hole pair. The valence band maximum and the conduction band minimum of direct semiconductors are located at the $k = 0$. In indirect semiconductors the extrema have different k-values; therefore, band-to-band recombination can only occur with the contribution of other quasi-particles that ensure k-conservation. These transitions are unsuitable for laser activity, because more partners are involved.

The probability that a state with the energy level E is occupied by an electron is expressed by the Fermi function $f(E,T)$.

$$f(E,T) = \frac{1}{\exp\left(\dfrac{E - E_F}{k_B T}\right) + 1} \tag{4.3}$$

At $T = 0$, the Fermi function is a step function which has a value of 1 (all electronic states filled) below the Fermi level energy E_F and a value of 0 (all states empty) for higher energies. In undoped semiconductors, the Fermi level is located between conduction and valence band edges. For higher temperatures T, the Fermi function is smeared out in the range $E_F \approx \pm k_B T$, with k_B being the Boltzmann constant.

The three types of radiative band-to-band transitions that can be found in semiconductors are sketched in Fig. 4.3. The first process is called spontaneous emis-

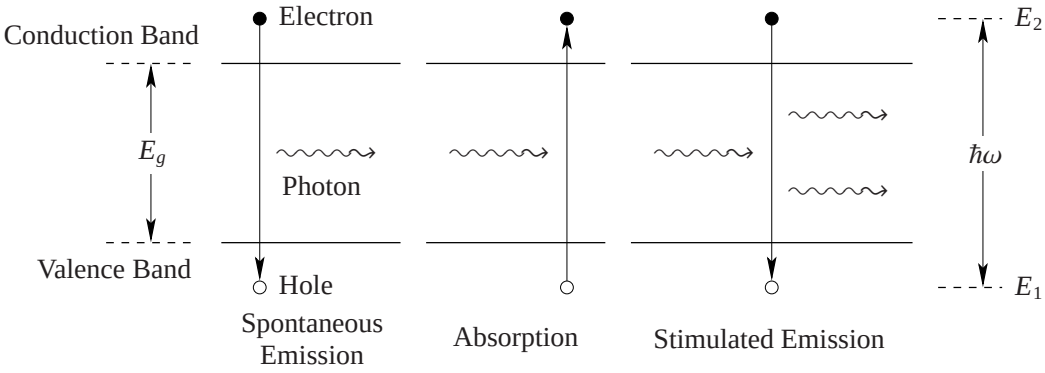

Figure 4.3: Radiative band-to-band transitions in semiconductors.

sion, where a recombination of an electron-hole pair leads to the emission of a photon. The emission of the photon is random in direction, phase, and time resulting in incoherent radiation. Since this process depends on the existence of an electron at E_2 and a hole at E_1 simultaneously, the transition rate for spontaneous emission $\mathcal{R}_{sp}$ is proportional to the product of the electron density at E_2 and the hole density at E_1. The electron density at the energy E_2 is the product of the density of electronic states $D(E_2)$ and the probability that they are occupied by electrons given by the Fermi function $f(E_2,T)$. The hole density at the energy

E_1 is the density of electronic states $D(E_1)$ multiplied with the probability of not being occupied by electrons $[1 - f(E_1, T)]$. So, the transition rate per volume for spontaneous emission of photons with fixed energy $\hbar\omega = E_2 - E_1$ can be written as

$$\mathcal{R}_{\mathrm{sp}} = A\, D(E_2) f(E_2, T)\, D(E_1)[1 - f(E_1, T)], \tag{4.4}$$

with A being the proportionality constant for spontaneous emission.

Absorption, also called stimulated absorption, is the second process illustrated in Fig. 4.3. A photon is absorbed and an electron-hole pair is generated. This is a three-particle process and the transition rate $\mathcal{R}_{12}$ therefore is proportional to the product of three particle densities: first, the density of nonoccupied states $D(E_2)[1 - f(E_2, T)]$ in the conduction band at the energy E_2, second, the density of states occupied by electrons $D(E_1) f(E_1, T)$ in the valence band at E_1, and third, the density of the photons $\rho(\hbar\omega)$ with energy $\hbar\omega = E_2 - E_1$.

$$\mathcal{R}_{12} = B_{12}\, \rho(\hbar\omega)\, D(E_1) f(E_1, T)\, D(E_2)[1 - f(E_2, T)]. \tag{4.5}$$

B_{12} is a proportionality constant for stimulated absorption.

The third process is stimulated emission. Recombination of an electron-hole pair is stimulated by a photon and a second photon is generated simultaneously which has the same direction and phase as the first photon. This process can be used to amplify optical radiation, since the photons are emitted into the optical mode of the stimulating photon resulting in coherent radiation. Light sources based on this emission process are, e.g., lasers, which is an abbreviation of light amplification by stimulated emission of radiation. Analogous to the stimulated absorption (4.5), the transition rate $\mathcal{R}_{12}$ for stimulated emission can be described as

$$\mathcal{R}_{21} = B_{21}\, \rho(\hbar\omega)\, D(E_2) f(E_2, T)\, D(E_1)[1 - f(E_1, T)], \tag{4.6}$$

with B_{12} being the proportionality constant for stimulated emission.

When the semiconductor is in thermal equilibrium with the photons, no energy is transferred form the semiconductor to the optical radiation field; thus, absorption and emission must be balanced:

$$\mathcal{R}_{12} = \mathcal{R}_{21} + \mathcal{R}_{\mathrm{sp}}. \tag{4.7}$$

Using (4.4), (4.5), and (4.6) for the rates $\mathcal{R}_{\mathrm{sp}}$, $\mathcal{R}_{12}$, $\mathcal{R}_{21}$, respectively, and inserting the Fermi function $f(E, T)$ from (4.3) and the relation $\hbar\omega = E_2 - E_1$ (4.2), gives

$$\rho(\hbar\omega) = \frac{A}{B_{12} \exp\left(\frac{\hbar\omega}{k_B T}\right) - B_{21}}. \tag{4.8}$$

The spectral energy density $u(\nu)\, d\nu$ at the frequency ν in a medium with refractive index n for radiation in thermal equilibrium is given by Planck's formula for blackbody radiation.

$$u(\nu)\, d\nu = \frac{8\pi h n^3 \nu^3}{c^3} \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1}\, d\nu. \tag{4.9}$$

Dividing the energy density $u(\nu)$ by the photon energy $\hbar\omega$ yields the photon density $\rho(\hbar\omega)$. Additionally, the relations $\omega = 2\pi\nu$, $h = 2\pi\hbar$, and $d(\hbar\omega) = d(h\nu) = h\,d\nu$ have been used.

$$\rho(\hbar\omega)\,d(\hbar\omega) = \frac{n^3(\hbar\omega)^2}{\pi^2\hbar^3c^3}\,\frac{1}{\exp\left(\frac{\hbar\omega}{k_BT}\right)-1}\,d(\hbar\omega). \tag{4.10}$$

In the thermal equilibrium of the semiconductor material with the radiation field, the spectral photon density described in (4.8) must be identical with the photon density of the blackbody radiation described by (4.10). Comparing these equations gives

$$B_{12} = B_{21} = B, \qquad A = \frac{n^3}{\pi^3\hbar^3c^3}(\hbar\omega)^2B. \tag{4.11}$$

Now we address the nonequilibrium conditions. If a p-n junction is forward biased, electrons and holes are injected into the depletion region where they can either recombine or travel further to the other side of the junction and recombine there with the majority carriers. In the transition zone the carrier distribution cannot be described by a single equilibrium Fermi function (4.3). Separate quasi-Fermi functions are used for the electrons $f_c(E,T)$ and for holes $f_v(E,T)$.

$$f_c(E,T) = \frac{1}{\exp\left(\frac{E-E_{Fc}}{k_BT}\right)+1}, \quad f_v(E,T) = \frac{1}{\exp\left(\frac{E-E_{FvBig}}{k_BT}\right)+1}. \tag{4.12}$$

Here the different Fermi levels are employed for the carrier distribution in the conduction (E_{Fc}) and valence (E_{Fv}) bands. The nonequilibrium situation can be described by replacing $f(E_1,T) \to f_v(E_1,T)$ and $f(E_2,T) \to f_c(E_2,T)$.

To determinate whether an optical wave with quantum energy $\hbar\omega$ is absorbed or amplified by stimulated emission, we need to know the ratio of the corresponding rates $\mathcal{R}_{12}$ and $\mathcal{R}_{21}$:

$$\frac{\mathcal{R}_{12}}{\mathcal{R}_{21}} = \frac{f_v(E_1,T)[1-f_c(E_2,T)]}{f_c(E_2,T)[1-f_v(E_1,T)]} = \exp\left[\frac{\hbar\omega-(E_{Fc}-E_{Fv})}{k_BT}\right]. \tag{4.13}$$

The result does not depend on the specific density of states $D(E)$. In thermal equilibrium $E_{Fc} = E_{Fv} = E_F$, the exponent $[\hbar\omega/(k_BT)]$ is positive, the exponential function is larger than 1, and therefore the absorption rate $\mathcal{R}_{12}$ is always larger than the rate $\mathcal{R}_{21}$ of the stimulated emission. Amplification in semiconductors can be only achieved if the condition

$$E_{Fc} - E_{Fv} > \hbar\omega > E_g \tag{4.14}$$

is fulfilled. In this inversion state the exponential function is smaller than 1 and the rate of stimulated emission is larger than the absorption rate. Laser operations requires pumping which builds up and maintains a nonequilibrium carrier distribution in the semiconductor material. From (4.13) it is clear that a laser-active transition always shows absorption in thermal equilibrium. Although pumping

can also be provided by optical excitation of electron-hole pairs, one main advantage of semiconductor laser over other types of lasers is the fact that they can be easily pumped with electrical currents as a forward-biased semiconductor diode as shown in Fig. 4.4. For this reason, electrically pumped semiconductor lasers

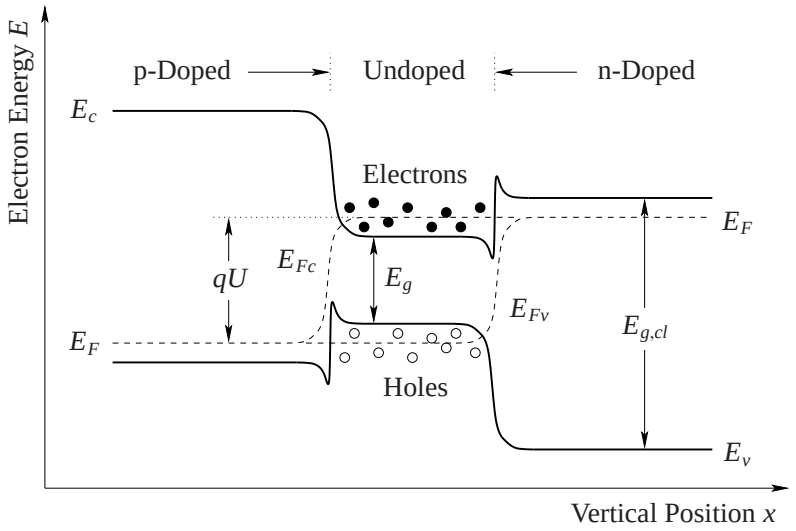

Figure 4.4: Forward-biased double-heterostructure p-i-n junction. Conduction E_c and valence band edges E_v are plotted as solid lines. The Fermi level energy E_F, represented by dashed lines, splits into quasi-Fermi levels E_{Fc} and E_{Fv} in the undoped transition region, where holes and electrons coexist when an external voltage U is applied. In this region, inversion is achieved since the quasi-Fermi levels are separated by more than E_g.

are called diode lasers.

Semiconductor-diode lasers use forward-biased double-hetero p-i-n structures to achieve carrier inversion. In this type of structure, an undoped semiconductor layer with a direct band gap is sandwiched between p-doped and n-doped material with a higher band gap. When the junction is forward biased, the quasi-Fermi levels E_{Fc} and E_{Fv} in the intrinsic layer are located inside the conduction and valence bands as illustrated in Fig. 4.4. This region acts as a laser-active layer which amplifies optical radiation by stimulated emission. Furthermore, the double heterostructure has two additional advantages.

- First, the carriers are confined between the double heterobarriers in the conduction and the valence bands and are therefore forced to recombine inside the intrinsic layer of the direct semiconductor material.

- Second, this layer sequence works like and optical waveguide since of most semiconductor–material systems, the low–band–gap layer in the middle of the structure has a higher refractive index.

The proportion of the carriers that recombines by stimulated and spontaneous emission can be determined using (4.4), (4.6), and (4.11).

$$\frac{\mathcal{R}_{21}}{\mathcal{R}_{\mathrm{sp}}} = \frac{B}{A}\rho(\hbar\omega) = \frac{\pi^2\hbar^3c^3}{n^3(\hbar\omega)^2}\rho(\hbar\omega). \tag{4.15}$$

From this equation one can see that a high photon density $\rho(\hbar\omega)$ in necessary to suppress spontaneous emission. Since the term $(\hbar\omega)^2$ is in the denominator of (4.15), a higher value of the photon density $\rho(\hbar\omega)$ is required for lasers with higher photon energy $(\hbar\omega)$ to achieve the same suppression of spontaneous emission. To obtain a high photon density in a semiconductor laser, optical waveguides are implemented to confine the photons in the laser-active region of the device. Furthermore, an optical resonator, mostly a Fabry-Perot resonator, is used to increase the photon density in the resonator cavity. A semiconductor laser can be regarded as an optical oscillator consisting of an optically amplifying medium and a resonator which provides optical feedback to the amplifier. Waveguides and resonators for high-power semiconductor lasers are discussed in more detail in the next subsections.

4.1.3 Basic elements of semiconductor diode lasers

There two constructing elements needed for a simple laser (87):

- a medium that provide an optical gain by stimulated emission;
- a resonator needed for the optical feedback.

Modern devices consist of at least two more vital elements (86):

- an optical waveguide that confine the photons in the active region:
- a lateral confinement of the injected carriers and photons that is required for operation in a fundamental mode.

The optical gain medium consists of an active undoped layer of a direct semiconductor material embedded between high-band-gap p- and n-doped regions. When this p-i-n junction is forward biased, electrons and holes are injected into the active region and optical gain by stimulated emission might be created. Furthermore, the double heterobarriers confine the carriers to a typical thickness of 100 nm in one or more QWs having typical thickness of 10 nm.

A dielectric optical waveguide consists of a core film with high refractive index embedded in cladding material with lower refracting index. The optical waveguide for a DH laser is shown in Fig. 4.5. The active film with band gap E_g, refractive index n_f, and thickness d is sandwiched between cladding layers with band gap $E_{g,cl}$ and refractive index n_{cl}. If the index step $\Delta n = n_f - n_{cl}$ and the core thickness d of the waveguide are small enough, only the fundamental mode with nearly Gaussian field distribution can propagate in the waveguide.

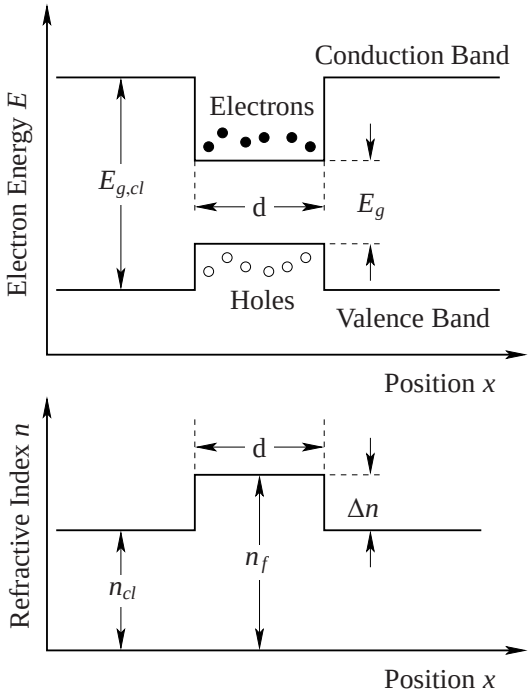

Figure 4.5: Confinement of the electronic carriers (electrons and holes) and the electric field (photons) using a double heterostructure in growth direction x of an edge-emitting diode laser. Plotted are the energy-band diagram $E(x)$ with conduction and valence bands (*top*), and the refractive-index profile $n(x)$ of the waveguide (*bottom*).

The optical wave traveling in the direction of the waveguide experiences an effective refraction index n_{eff} which is different from the refractive indices of core and cladding ($n_{cl} \leq n_{eff} \leq n_f$). Fig. 4.5 shows a structure where the sample layer provides the confinement of the carriers and the optical wave. In quantum-well lasers, so-called separate confinement structures are implemented, where the carriers are confined in the quantum well and the optical wave is confined in a separate dielectric-waveguide structure.

For high-power diode lasers, Fabry-Perot resonators are use. Figure 4.6 depicts this type of resonator consisting of two mirrors with distance L around a laser-active material having an optical waveguide with effective refractive index n_{eff} in a propagation direction normal to the mirror surfaces. The resonator pro-

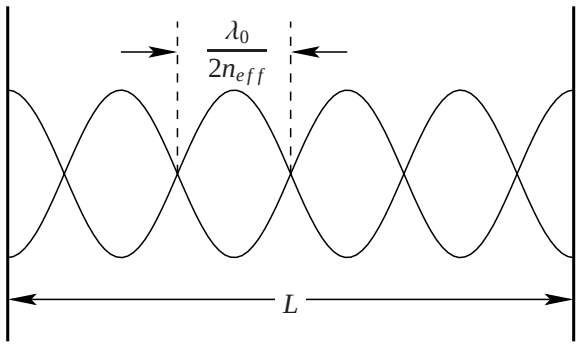

Figure 4.6: A standing wave having $m = 5$ nodes in a Fabry-Perot resonator with a cavity length L. The wave propagates in a waveguide with an effective refractive index n_{eff}. The distance between two nodes is $\lambda_0/(2n_{eff})$ with λ_0 being the vacuum wavelength.

vides feedback, when a standing wave develops between the mirrors.

$$L = m\,\frac{\lambda_0}{2n_{eff}}, \qquad m = 1,\ 2,\ 3,\ \ldots\ . \tag{4.16}$$

m is the number of nodes of the standing wave, the order number of the longitudinal mode, and λ_0 is the vacuum wavelength.

The lateral confinement, perpendicular to the growth direction, is implemented either by gain- or index-guiding. All structures investigated in this work are gain-guided structures.

4.1.4 Optical gain and threshold condition

The intensity of a planar optical wave exponentially decreases when its going through an absorbing medium.

$$J(z) = J_0 \exp(-\alpha z). \tag{4.17}$$

J_0 is initial intensity and α is the absorption coefficient. In the laser-active semiconductor materials, an amplification of the optical waves is achieved. Then, the exponential increase in intensity can be thought as a negative value of α which is called optical gain $g = -\alpha$. In an optical waveguide, only a part of the intensity is located in the core of the waveguide. One has to distinguish between the gain of the active material itself, called the material gain g, and the significantly lower gain of the optical mode, called the modal gain g_{modal}.

In Fig. 4.7, the typical behavior of the material gain is shown. For regular

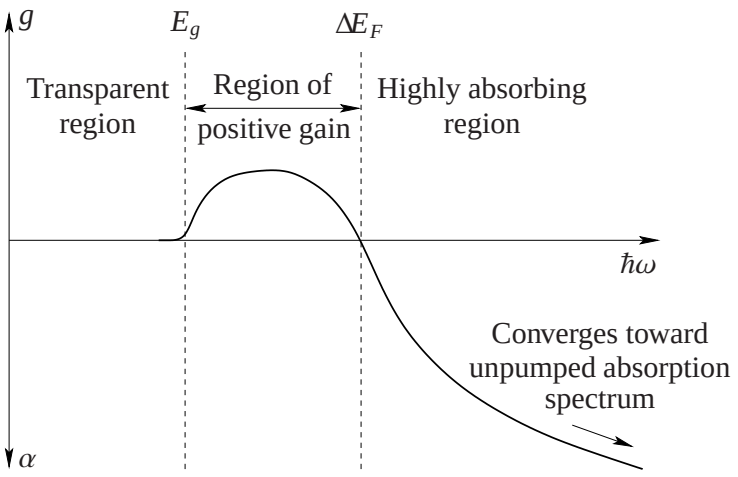

Figure 4.7: Gain spectra in bulk matrial.

electron-hole recombination the maximum gain is observed at photon energies which are slightly higher than the band gap energy. Fig. 4.8 illustrates the optical-intensity pattern $J(x)$ of the fundamental optical mode in a double heterostructure edge-emitting laser having an active-layer thickness d. The relation between modal gain g_{modal} and material gain g is expressed by defining a confinement factor

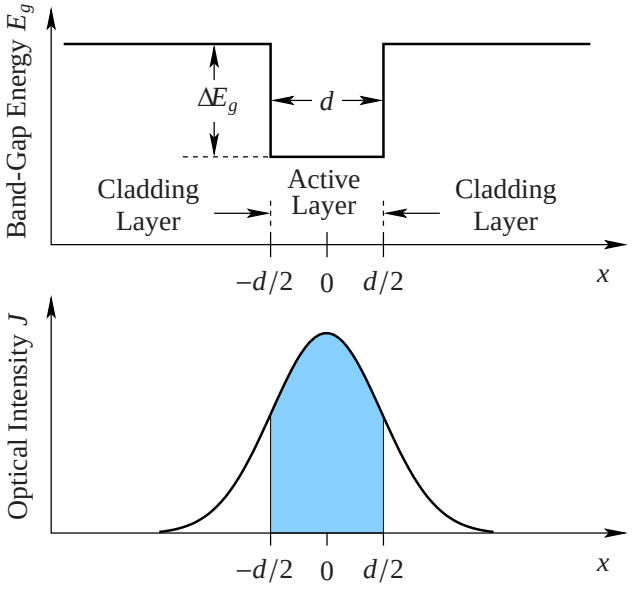

Figure 4.8: The confinement factor Γ is defined by the overlap of the intensity pattern J of the optical mode with the active region. The illustration shows the intensity distribution $J(x)$ in the vertical direction x of a fundamental optical mode with nearly Gaussian shape for an edge-emitting laser with and active-layer thickness d.

Γ which depends on the overlap of the optical-mode pattern with the gain region of the laser.

$$g_{\text{modal}} = \Gamma g, \qquad \Gamma = \frac{\int_{-d/2}^{d/2} J(x)\,dx}{\int_{-\infty}^{+\infty} J(x)\,dx}. \tag{4.18}$$

In double heterostructures with active-layer thickness of 50-300 nm, the confinement factor Γ has values in the range 10-70 %. If the active layer consists of a quantum well with a typical thickness around 10 nm, confinement factors of a few percent are obtained.

For a mode traveling along the optical waveguide, the intensity-absorption coefficient α is usually split into two parts, one describing the modal absorption α_{i} and the other describing the modal gain $g_{\text{modal}} = \Gamma g$ which depends on the density of the injected carriers.

$$\alpha = \alpha_{\text{i}} - \Gamma g. \tag{4.19}$$

The intrinsic modal absorption in caused by scattering of the optical mode at defects or rough interfaces and by free-carrier absorption. Whereas scattering is extremely low for semiconductor diode lasers with good crystalline quality, free-carrier absorption cannot be avoided since part of the optical-mode pattern overlaps with the p- and n-doped cladding regions. Also the non-equilibrium carriers in the active layer contribute to free-carrier absorption When the modal gain Γg is larger than the modal loss α_{i}, the propagation optical mode is amplified.

In a laser, the optical waveguide is combined with a Fabry-Perot resonator consisting two mirrors with reflectivities R_1 and R_2. Part of the optical intensity leaves the cavity through this mirrors as a laser output beam. As shown in Fig.

4.9, the intensity of the of the optical mode after a roundtrip in the cavity is given

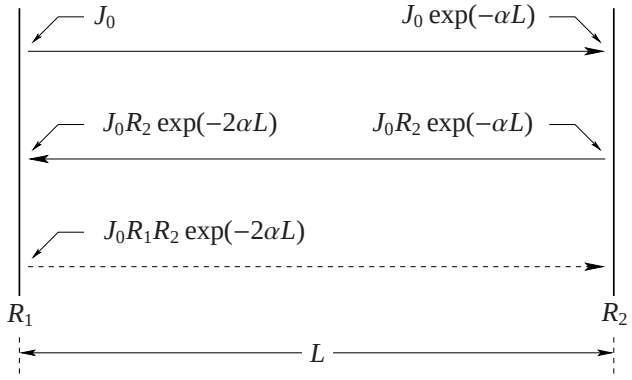

Figure 4.9: Intensity of an optical wave during a roundtrip in a Fabry-Perot resonator with cavity length L and mirror reflectivities R_1 and R_2.

by

$$J_{\mathrm{rt}} = J_0 R_1 R_2 \exp[2(\Gamma g - \alpha_{\mathrm{i}})L]. \tag{4.20}$$

Lasing occurs when the gain provided to the optical mode compensates the intrinsic absorption and the mirror losses for a roundtrip. The minimum gain g where the device starts lasing operation is called the threshold gain g_{th}. In this case, the intensity J_{rt} after a roundtrip in the cavity again has its initial value J_0.

$$\begin{aligned} J_{\mathrm{rt}} &= J_0, \\ 1 &= R_1 R_2 \exp[2(\Gamma g - \alpha_{\mathrm{i}})L], \\ \Gamma g_{\mathrm{th}} &= \alpha_{\mathrm{i}} + \frac{1}{2L} \ln\left(\frac{1}{R_1 R_2}\right) = \alpha_{\mathrm{i}} + \alpha_{\mathrm{mirror}}. \end{aligned} \tag{4.21}$$

At laser threshold, the modal gain Γg_{th} is the sum of the two terms in (4.21), the intrinsic absorption α_{i} and the mirror losses α_{mirror}. The mirror losses depend on the cavity length L and the mirror reflectivities R_1 and R_2.

4.1.5 Quantum well structures

In double heterostructures, the typical thickness of the active layer is d = 50-300 nm, resulting in a confinement factor Γ (overlap of the optical-mode pattern with the gain region of the laser (86)) in the range of 10-70 %. The density of electronic states $D(E)$ increases with the square root of the energy at the band edge ($D(E) \propto \sqrt{E - E_g}$). If the thickness of the active layer is shrunk to values of 5-10 nm, the electronic wavefunction of this QW show quantization in the vertical direction x resulting in discrete energy levels. In this case, the density of electronic states $D(E)$ increases in steps which are located at the electronic energy levels of the QW.

Due to the small electronically active volume of a QW laser, a reduced threshold current can be obtained. Additionally, the material gain is higher and the spectral shift of the gain curve due to the band-filling effect is much smaller, because of the higher carrier density and its narrower energetic distribution. One

of the advantages in QW structures, is the possibility to introduce mechanical strain that can be used for band gap tailoring. The useable wavelength range of a particular material system can also be extended beyond the range reachable with alloying technique only. So, the incorporation of In instead of Ga into a thin GaAs QW layer results in a compressively strained QW and the accessible wavelength now ranges from 870 nm for bulk GaAs into the long-wavelength region up to approximately 1100 nm.

The typical spectral gain behavior for a QW is shown in Fig. 4.10 (see Fig. 4.7 for comparison with bulk material gain).

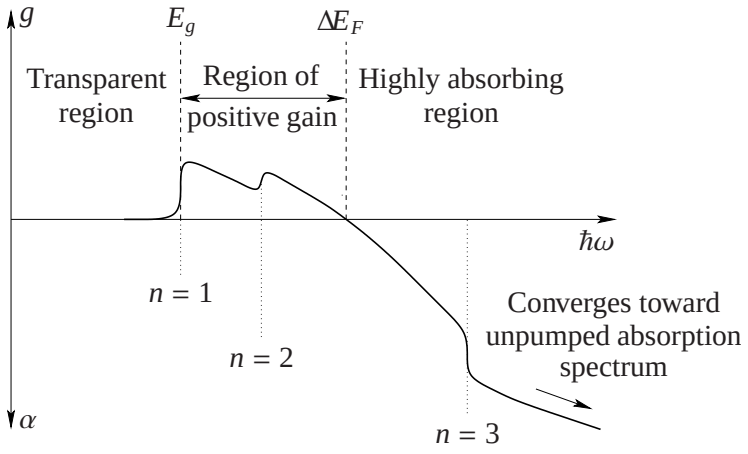

Figure 4.10: Schematic diagram of a gain spectrum of a QW material.

The product of QW film thickness and strain must be below a critical value. Above this value, the film experiences relaxation, which is associated with a high number of native defects. In the long-wavelength range, any kind of strain is beneficial due to the reduced inter-valence-band absorption and Auger recombination. Also a significantly improved reliability has been observed for compressively strained QW lasers. This is one of the reasons why the strained QWs are the rule rather than the exception in state of the art diode lasers.

Since QWs are very thin, the confinement of the optical mode is poor. This can be overcome by Separate-Confinement Heterostructure (SCH) where the confinement of the optical mode is provided by a separate waveguide structure. Two examples of such vertical structures are shown in Fig. 4.11. If the waveguide includes a graded refractive-index profile, the structure is called a GRaded-INdex Separate-Confinement Heterostructure (GRINSCH).

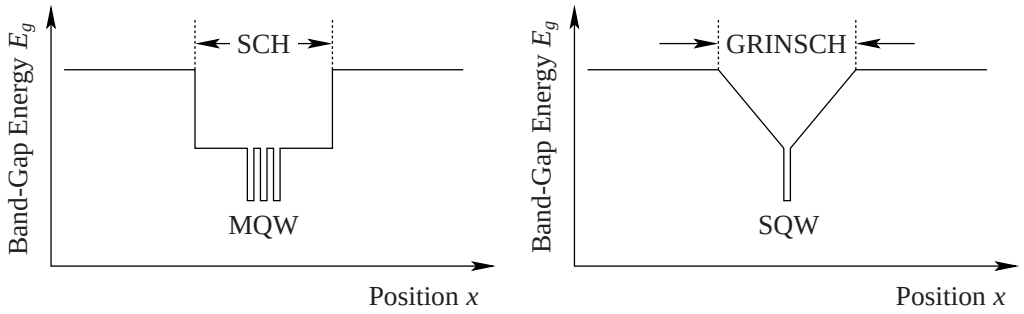

Figure 4.11: A Multi-QW SCH with three WQs (left side). A Single-QW GRIN-SCH (right side). The left structure is also called step-index structure.

4.2 Waveguiding modes and the method of the fake waveguide

4.2.1 Eigenmodes of semiconductor laser waveguides

In this subsection we will address the eigenmodes of a semiconductor laser waveguide. For the simplest case of a planar dielectric waveguide it is possible to find an analytical solution. Knowing it one can use perturbation theory for the more complicate case where QWs are situated inside of the waveguide.

Such an approach is working very well for step-index waveguides that is a slab of dielectric material surrounded by media of lower refractive indices. The light is guided inside the slab by total internal reflection. In thin-film devices the slab is called the "waveguide" and the upper and lower media are called the "cladding layers". The inner medium and outer media may also be called the "core" and the "cladding" of a waveguide, respectively.

This way of solving the eigenmode problem has significant difficulties for more complicated systems as graded-index waveguides, coupled and weakly coupled waveguides or for systems where no analytical solution is found.

In order to approach the problem we use the method of fake waveguide. It is fully numerical method based on first order perturbation theory. The main advantage of this method is that it can describe optical systems with no limit on the degree of complexity. The idea is to treat not only QWs inside of waveguide as a perturbation but whole waveguide system as perturbation to some ideal system.

As the ideal system we choose the planar waveguide resonator formed by two ideal mirrors with an uniform dielectric medium between them. The eigenmode equation of this system has an analytical solution (96). Now we can form any type of waveguide structure inside of the resonator by varying the refraction index between the mirrors. Furthermore we adjust the mirrors to a position distance where they do not affect the newly formed structure. This variation will be treated as perturbation. Usually, in this way one can get a huge amount of "parasitic" modes that are located outside the waveguide. But they are artifacts created by the

presence of the ideal mirrors. The mirrors do not exist in the real system, but are used for analytical/numerical convenience.

We start from an analysis of the ideal system shown in Fig. 4.12. Light rays

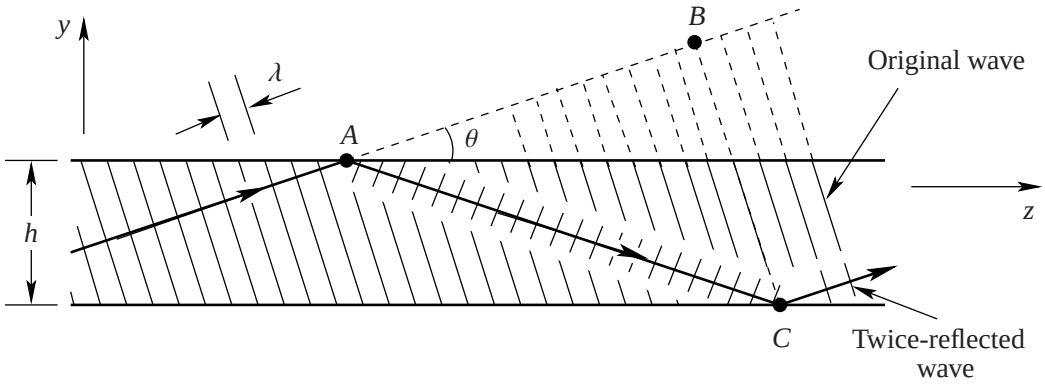

Figure 4.12: Homogeneous metal waveguide and wave propagating in z-direction.

arriving at angles θ with the z axis in the y–z plane, undergo multiple reflections at the mirrors. They travel in the z-direction by bouncing between mirrors without loss of power.

Assuming that the field in the slab is of the form of a monochromatic plane wave with a wavelength $\lambda = \lambda_0/n$, wavenumber $k = nk_0$ and phase velocity $c = c_0/n$, where n is the refractive index of the medium between the mirrors. The wave is polarized in the x direction and its wavevector lies in the y–z plane at an angle θ with z axis. Since the electric field is parallel to the mirrors, each reflection produces a phase shift π with no change in the amplitude and polarization. The π phase shift ensures that the sum of the waves and its own reflection vanishes so that the total field is zero at the mirrors. At each point within the waveguide we have transverse electromagnetic (TEM) plane waves traveling into the upward direction at an angle θ and others traveling into the downward direction at an angle $-\theta$; all waves are polarized in x direction.

Now we impose a self-consistency condition by requiring that as the wave reflects twice, it reproduces itself, so that we have only two distinct plane waves. Fields that satisfy this conditions are called eigenmodes or simply modes of the waveguide. *Modes are fields that maintain the same transverse distribution and polarization at all distances along the waveguide axes.* We shall see that self-consistency guarantees this shape invariance. The phase shift encountered by the original wave in traveling form A to B (Fig. 4.12) must be equal to, or different by an integer multiple of 2π, from that encountered when the wave reflects, travels from A to C, and reflects once more. Accounting for a phase shift of π at each reflection, we have $2\pi\overline{AC}/\lambda - 2\pi - 2\pi\overline{AB}/\lambda = 2\pi q$, where $q = 0, 1, 2, \ldots$. Since $\overline{AC} - \overline{AB} = 2h\sin\theta$, where h is the distance between the mirrors, $2\pi(2h\sin\theta)/\lambda = 2\pi(q+1)$, and

$$\frac{2\pi}{\lambda} 2h\sin\theta = 2\pi m, \qquad m = 1, 2, \ldots, \tag{4.22}$$

where $m = q + 1$. The self-consistency condition is therefore satisfied only for certain bounce angles $\theta = \theta_m$ satisfying

$$\sin\theta_m = m\frac{\lambda}{2h}, \qquad m = 1, 2, \ldots. \tag{4.23}$$

Each integer m corresponds to a bounce angle θ_m, and the corresponding fields are called the m^{th} mode. The $m = 1$ mode has the smallest angle $\theta_1 = \sin^{-1}(\lambda/2h)$; modes with larger m are composed of more oblique plane-wave components.

When the self-consistency condition is satisfied, the phases of the upward and downward plane waves at points on the z axis differ by half the round-trip phase shift $q\pi$, $q = 0, 1, \ldots$, or $(m-1)\pi$, $m = 1, 2, \ldots$, so that they add for odd m and substract for even m.

Since the y component of the propagation constant is $k_y = nk_0 \sin\theta$, it is quantized to the values $k_{ym} = nk_0 \sin\theta_m = (2\pi/\lambda)\sin\theta_m$. Using (4.23), we obtain

$$k_{ym} = m\frac{\pi}{h}, \qquad m = 1, 2, 3, \ldots, \tag{4.24}$$

so that the k_{ym} are spaced by π/h. Equation (4.24) states that the phase shift encountered when a wave travels a distance $2h$ (one round trip) in the y direction, with propagation constant k_{ym}, must be a multiple of 2π.

The guided wave is composed of two distinct plane waves traveling at angles $\pm\theta$ with the z axis in the y–z plane. Their wavevectors have components $(0, k_y, k_z)$ and $(0, -k_y, k_z)$. Their sum or difference therefore varies with z as $\exp(-ik_z z)$, so that the propagation constant of the guided wave is $\beta = k_z = k\cos\theta$. Thus β is quantized to the values $\beta_m = k\cos\theta_m$, from which $\beta_m^2 = k^2(1 - \sin^2\theta_m)$. Using (4.23), we obtain

$$\beta_m^2 = k^2 - \frac{m^2\pi^2}{h^2}. \tag{4.25}$$

Higher-order (more oblique) modes travel with smaller propagation constants.

The complex amplitude of the total field in the waveguide is the superposition of the reflecting TEM plane waves. If $A_m \exp(-ik_{ym}y - i\beta_m z)$ is the upward wave, then $e^{i(m-1)\pi}A_m \exp(+ik_{ym}y - i\beta_m z)$ must be the downward wave [at $y = 0$, the two waves differ by a phase shift $(m-1)\pi$]. Therefore there are symmetric modes, for which the two plane-wave components are added, and antisymmetric modes, for which they are subtracted. The total field turns out to be $E_x(y, z) = 2A_m \cos(k_{ym}y)\exp(-i\beta_m z)$ for odd modes and $2iA_m \sin(k_{ym}y)\exp(-i\beta_m z)$ for even modes.

Using (4.24) we write the complex amplitude of the electric field in the form

$$E_x(y, z) = a_m\, u_m(y)\, \exp(-i\beta_m z), \tag{4.26}$$

where

$$u_m(y) = \begin{cases} \sqrt{\dfrac{2}{h}}\cos\dfrac{\pi m}{h}y, & m = 1, 3, 5, \ldots \\ \sqrt{\dfrac{2}{h}}\sin\dfrac{\pi m}{h}y, & m = 2, 4, 6, \ldots, \end{cases} \tag{4.27}$$

and $a_m = \sqrt{2h}A_m$ and $i\sqrt{2h}A_m$, for odd and even m, respectively. The funcitons $u_m(y)$ have been normalized

$$\int_0^h u_m^2(y)\,dy = 1. \tag{4.28}$$

Thus a_m is the amplitude of mode m. It can be shown that the functions $u_m(y)$ also satisfy

$$\int_0^h u_m(y)\,u_l(y)\,dy = 0, \qquad l \neq m, \tag{4.29}$$

i.e., they are orthogonal in the $[0, h]$ interval.

The transverse distributions $u_m(y)$ are plotted in Fig. 4.13. Each mode can be

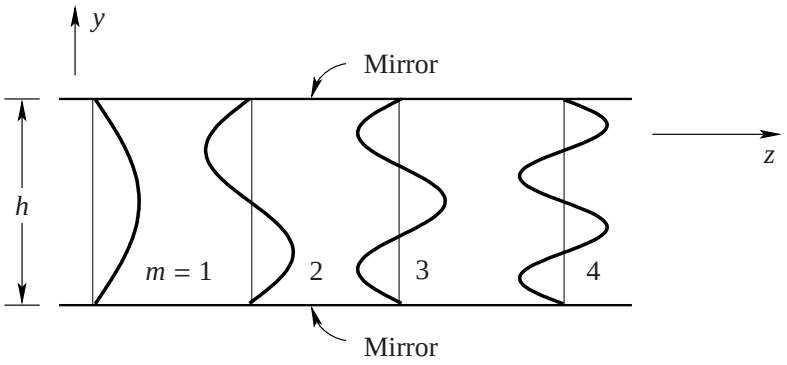

Figure 4.13: Field distributions of the modes of a planar-mirror waveguide.

considered a standing wave in y direction, traveling in z direction. Modes of large m vary in the transverse plane at larger k_y and travel with a smaller propagation constant β. The field vanishes at $y = 0$ and $y = h$ for all modes, so that the boundary conditions at the surface of the mirrors are always satisfied.

Since we assumed that the bouncing TEM plane wave is polarized in x direction, the total electric field is also in the x direction and the guided wave is a transverse electric (TE) wave. Transverse magnetic (TM) waves may be treated similarly.

Since $\sin\theta_m = m\lambda$, $m = 1, 2, \ldots$ and for $\sin\theta_m < 1$, the maximum allowed value of m is the greatest integer smaller than $(\lambda/2h)^{-1}$,

$$M \doteq \frac{2h}{\lambda}. \tag{4.30}$$

Thus M is the number of modes of the waveguide. Light can be transmitted through the waveguide in one, two, or many modes. The actual number of modes that carry optical power depends on the source of excitation, but the maximum number is M.

The number of modes increases with increasing ratio of the mirror separation to the wavelength. If $2h/\lambda \leq 1$, $M = 0$, indicating that the self-consistency condition cannot be met and the waveguide cannot support any modes. The wavelength

$\lambda_{max} = 2h$ is called the *cutoff wavelength* of the waveguide. It is the longest wavelength that can be guided by the structure. It corresponds to the *cutoff frequency* $\nu_{min} = c/2h$, the lowest frequency of light that can be guided by the waveguide. If $1 < 2h/\lambda \leq 2$ (i.e., $h \leq \lambda \leq 2h$), only one mode is allowed. The structure is a *single-mode waveguide*. Equation (4.30) can also be written in terms of the frequency ν, $M = \nu/(c/2h)$, so that the number of modes increases with the frequency ν.

A pulse of light (wavepacket) of angular frequency centered at ω and propagation constant β travels with a velocity $v = d\omega/d\beta$, known as the group velocity. The propagation constant of mode m is given by (4.25) from which $\beta_m^2 = (\omega/c)^2 - m^2\pi^2/h^2$, which is an explicit relation between β_m and ω known as the *dispersion relation*. Taking the derivative and assuming that c is independent of ω (i.e., ignoring dispersion in the waveguide material), we obtain $2\beta_m d\beta_m/d\omega = 2\omega/c^2$, so that $d\omega/d\beta_m = c^2\beta_m/\omega = c^2 k\cos\theta_m/\omega = c\cos\theta_m$, from which the group velocity of mode m is

$$v_m = c\cos\theta_m. \tag{4.31}$$

Thus different modes have different group velocities. More oblique modes travel with a smaller group velocity since they are delayed by the longer path of the zigzagging process.

TM modes (magnetic field in x direction) can also be supported by the mirror waveguide. They can be described by means of a TEM plane wave with the magnetic field in the x direction, traveling at an angle θ and being reflected from the two mirrors. The electric field complex amplitude then has components in the y and z directions. Since the z component is parallel to the mirror, it must behave like the x component of the TE mode (i.e., undergoes a phase shift π at each reflection and vanishes at the mirror). When the self-consistency condition is applied to this component the result is mathematically identical to that of the TE case. The angles θ, the transverse wavevector components k_y, and the propagation constants β of the TM modes associated with this component are identical to those of the TE modes. There are $M \doteq 2h/\lambda$ TM modes (and a total of $2M$ modes) supported by the waveguide.

As previously, the z component of the electric field complex amplitude of mode m is the sum of an upward plane wave $A_m \exp(-ik_{ym}y)\exp(-i\beta_m z)$ and a downward plane wave $e^{i(m-1)\pi}A_m \exp(ik_{ym}y)\exp(-i\beta_m z)$, with equal amplitudes and phase shifts $(m-1)\pi$, so that

$$E_z(y,z) = \begin{cases} a_m\sqrt{\dfrac{2}{h}}\cos\dfrac{\pi m y}{h}\exp(-i\beta_m z), & m = 1,3,5,\ldots \\ a_m\sqrt{\dfrac{2}{h}}\sin\dfrac{\pi m y}{h}\exp(-i\beta_m z), & m = 2,4,6,\ldots, \end{cases} \tag{4.32}$$

where $a_m = \sqrt{2h}A_m$ and $i\sqrt{2h}A_m$ for odd and even m, respectively. Since the electric field vector of a TEM plane wave is normal to its direction of propagation,

it makes an angle $\pi/2 + \theta_m$ with the z axis for the upward wave, and $\pi/2 - \theta_m$ for the downward wave.

The y component of the electric field of these waves are

$$A_m \cot\theta_m \exp(-ik_{ym}y)\exp(-i\beta_m z)$$

and

$$e^{im\pi} A_m \cot\theta_m \exp(ik_{ym}y)\exp(-i\beta_m z),$$

so that

$$E_y(y,z) = \begin{cases} a_m \sqrt{\dfrac{2}{h}} \cot\theta_m \cos\dfrac{\pi m y}{h} \exp(-i\beta_m z), & m = 1,3,5,\ldots \\ a_m \sqrt{\dfrac{2}{h}} \cot\theta_m \sin\dfrac{\pi m y}{h} \exp(-i\beta_m z), & m = 2,4,6,\ldots. \end{cases} \tag{4.33}$$

Satisfaction of the boundary conditions is assured because $E_z(y,z)$ vanishes at the mirrors. The magnetic field component $H_x(y,z)$ may be similarly determined by the ratio of the electric to the magnetic fields of a TEM wave is the impedance of the medium η. The resultant fields $E_y(y,z)$, $E_z(y,z)$, and $H_x(y,z)$ do, of course, satisfy Maxwell's equations.

4.2.2 First order perturbation theory for waveguides

In this subsection we introduce the perturbation approach (97) for solving complex problems. In order to use it we generally need knowledge about the eigenmodes of a simple waveguide configuration. Then, the idea is to express the solution of some perturbed or more complex configuration in terms of this original basis set of eigenmodes (see appendix A).

For mode m the electric field can be written as

$$\vec{E_m}(x,y,z,t) = \hat{\mathbf{e}}_j E_{0m} u_m(x,y) e^{i(\omega t - \beta_m z)}, \tag{4.34}$$

where $\hat{\mathbf{e}}_j$ is the unit vector along the j^{th} coordinate (giving the polarization direction), E_{0m} gives the magnitude of the field, and u_m is the normalized transverse mode shape for mode m. For convenience, we can also combine the polarization into u_m, so that $\mathbf{u}_m = \hat{\mathbf{e}}_j u_m$. Thus, making use of the orthogonality between eigenmodes, we have

$$\int u_m^* \cdot u_n dA = \delta_{mn}, \tag{4.35}$$

where δ_{mn} is the Kronecker delta function. Orthogonal modes of a uniform waveguide do not interact. It is also worth mentioning that in the regular way of solving a waveguide problem, one finds that the eigenfunctions u_m provide a complete set.

Given that the u_m form a complete set, we can express an arbitrary field in the vicinity of a waveguide by a normal-mode expansion of all of the waveguide eigenmodes (including unguided radiation modes). Thus,

$$\vec{E}(x,y,z,t) = \sum_m \vec{E}_m(x,y,z,t), \tag{4.36}$$

where the amplitudes of the various terms in the summation are given by E_{0m} in Eq. (4.34). Both Eqs. (4.34) and (4.36) are solutions to the wave equation,

$$\nabla^2\vec{E} + \varepsilon(x,y,z)k_0^2\vec{E} = 0, \tag{4.37}$$

where $\varepsilon(x,y,z)$ is the relative dielectric constant, and k_0 is the free-space propagation constant for the medium of interest. For a single mode of a waveguide, we can also use Eq. (4.34) in (4.37) to obtain a wave equation for the transverse mode profile, u:

$$\nabla_T^2 u + [\varepsilon(x,y,z)k_0^2 - \beta^2]u = 0, \tag{4.38}$$

where we have used $\nabla^2\vec{E} = \partial^2\vec{E}/\partial z^2 + \nabla_T^2\vec{E}$.

Although the higher order modes of even a simple waveguide may be complicated, it is fortunate that we need only to know the details of not more than two modes for all of the discussion to follow in this subsection. In fact, only one mode must be characterized for many considerations.

Many real waveguide structures involve a slight perturbation from a mathematically more simple structure, for which the eigenmode shapes, u_m, and propagation constants, β_m, are known. The perturbation can usually be expressed in terms of a change in relative dielectric constant, $\Delta\varepsilon$, which is generally complex.

If we replace ε by $\varepsilon + \Delta\varepsilon$ in Eq. (4.37), we must assume that Eq. (4.36) must be used to represent the perturbed field. First we consider sufficiently weak perturbations where scattering to other modes can be neglected. This might occur in a single-mode waveguide. Thus, in response to $\varepsilon \longrightarrow \varepsilon + \Delta\varepsilon$, let $\beta \longrightarrow \beta + \Delta\beta$, and $u \longrightarrow u + \Delta u$ in Eq. (4.38). That is,

$$\nabla_T^2(u + \Delta u) + [(\varepsilon + \Delta\varepsilon)k_0^2 - (\beta + \Delta\beta)^2](u + \Delta u) = 0. \tag{4.39}$$

Multiplying this out, and dropping the unperturbed transverse wave equation, which equals zero, and the second-order perturbation terms, we arrive at

$$\nabla_T^2\Delta u + \varepsilon k_0^2\Delta u + \Delta\varepsilon k_0^2 u - 2\beta\Delta\beta u - \beta^2\Delta u = 0. \tag{4.40}$$

Now we multiply by the complex conjugate of the transverse mode and integrate over the cross section. Because of modal orthogonality this helps to remove most of the unwanted terms and we obtain:

$$\begin{aligned} 2\beta\Delta\beta\int |u|^2 dA \;\; &= \;\; \int \Delta\varepsilon k_0^2 |u| dA + \\ &\quad + \int [(\nabla_T^2\Delta u)u^* + \varepsilon k_0^2\Delta u u^* - \beta\Delta u u^*]dA. \end{aligned} \tag{4.41}$$

The second term on the right is negligible, provided that both Δu and u vanish at infinity and that the unperturbed ε and β are mostly real. This indicates that the small change in the transverse mode shape, Δu, due to $\Delta\varepsilon$ will have *no effect* on the propagation constant to first order. Thus, we fortunately do not have to worry about how the changing mode shape will change the averaging, at least for

this first-order approximation, and we only need to know the original unperturbed transverse mode. The same conclusions we obtain from a more general quantum-mechanical consideration of the first order perturbation made in appendix A (Eq. (A.21)).

Then, solving for $\Delta\beta$, we have the desired perturbation formula

$$\Delta\beta = \frac{\int \Delta\varepsilon k_0^2 |u|^2 dA}{2\beta \int |u|^2 dA}. \tag{4.42}$$

If u is normalized according to Eq. (4.35), the denominator integral is just unity. Usually, the index perturbation is limited in lateral extend, and it may even be constant over some range. In these cases the integration is easily performed. It is important to keep in mind that all quantities, except for the actual perturbation, $\Delta\varepsilon$, are for the original unperturbed problem.

In the similar way we can develop an equation for the mode shape change caused by the perturbation. But we will just adopt results from appendix A (Eq. (A.27):

$$\vec{E_x} = \vec{E_{x_0}} + \sum_{p \neq n} \frac{\Delta\varepsilon \int_0^h E_{x_p} E_{E_{x_n}} dy}{\beta_p^2 - \beta_n^2}. \tag{4.43}$$

Now knowing the analytical solution for the mode profile in a metallic planar waveguide (4.27) we can write the desired equation for first order perturbed modes:

$$u_m(y) = \begin{cases} \dfrac{1}{h}\left[\dfrac{\sin \dfrac{\pi(m-p)y}{h}}{\dfrac{\pi(m-p)y}{h}} - (-1)^m \cdot \dfrac{\sin \dfrac{\pi(m+p)y}{h}}{\dfrac{\pi(m+p)y}{h}} \right], & m = 1, 3, 5, \ldots \\ \dfrac{1}{h}\left[\dfrac{-\cos \dfrac{\pi(m+p)y}{h}}{\dfrac{\pi(m+p)y}{h}} - (-1)^m \cdot \dfrac{\cos \dfrac{\pi(m-p)y}{h}}{\dfrac{\pi(m-p)y}{h}} \right], & m = 2, 4, 6, \ldots \end{cases} \tag{4.44}$$

Using (4.42), (4.43) and (4.44) we able to calculate the real mode profile and propagation constants for real laser diodes. The way is simply to filter out "artificial" modes that have a maxima outside the real waveguide structure. What will remain are the wanted solutions.

4.3 Near-field photocurrent imaging of the semiconductor laser diodes with a single waveguide

4.3.1 Waveguide mapping problem

Optical waveguides are key components of modern semiconductor diode lasers. They confine the optical mode inside the laser cavity to dimensions in the micron or even submicron range and thus ensures a sufficient overlap of the optical mode and the gain medium, e.g., a double quantum well (DQW) layer inside the waveguide. While the theoretical description of the waveguide mode structure by solving Helmholtz equation is well established (98; 99) (see section 4.2), it is generally quite complicated to experimentally image the optical mode structure directly. From an application point of view, such imaging is of considerable interest as it would allow to characterize the effects of waveguide imperfections (100; 101) or dopant profiles on the waveguide mode structure. This requires optical techniques with subwavelength spatial resolution, such as NSOM (42; 45), that overcome the diffraction-limited resolution of FF microscopy by using evanescent fields in the vicinity of nanometer-sized objects, such as apertures in metal-coated fiber probes (45) or sharp metal tips (102) (see section 2.4 and chapter 3). Information on transversal optical mode profiles of laser diodes was obtained by spatially resolved collection of the laser emission (103; 104). These techniques are limited to those modes that contribute to the laser emission. This limitation can be overcome and the mode structure of both the fundamental and of higher waveguide modes can be imaged by monitoring the position-dependent coupling of a localized light source into these modes using the recently introduced technique of NF photocurrent spectroscopy (105; 106; 107).

In this section, we demonstrate the potential of this technique for direct mapping of the optical mode profiles in waveguides by experimentally and theoretically studying NF photocurrent images of high-power laser diodes of different waveguide designs.

4.3.2 Samples under investigation

In our experiments, we investigate asymmetrically coated (In)AlGaAs/GaAlAs/GaAs high power diode lasers with different waveguide designs. All samples were grown by MOCVD method (see subsection 4.1.1) and consist of an active region of two 8 nm (In)AlGaAs QW inside an $Al_{0.3}Ga_{0.7}As$ step-index (SIN) waveguide that is clad by two 1.5 μm wide $Al_{0.6}Ga_{0.4}As$ layers. We compare structures with (i) the DQW centered inside a 440 nm wide SIN waveguide (diode A), (ii) a 1000 nm SIN waveguide (DQW symmetrically centered, diode B), and (iii) a 1000 nm SIN region with a DQW that is off-centered by 120 nm (diode C). All samples are antireflection (AR) coated with an ~120-nm-thick Al_2O_3-layer. The photon energy of the laser emission is 1.53 eV (λ = 808 nm), for all three diodes.

4.3.3 Direct mapping of the optical mode profiles

In the near-field photocurrent (NPC) experiments as shown in Fig. 4.14, the laser

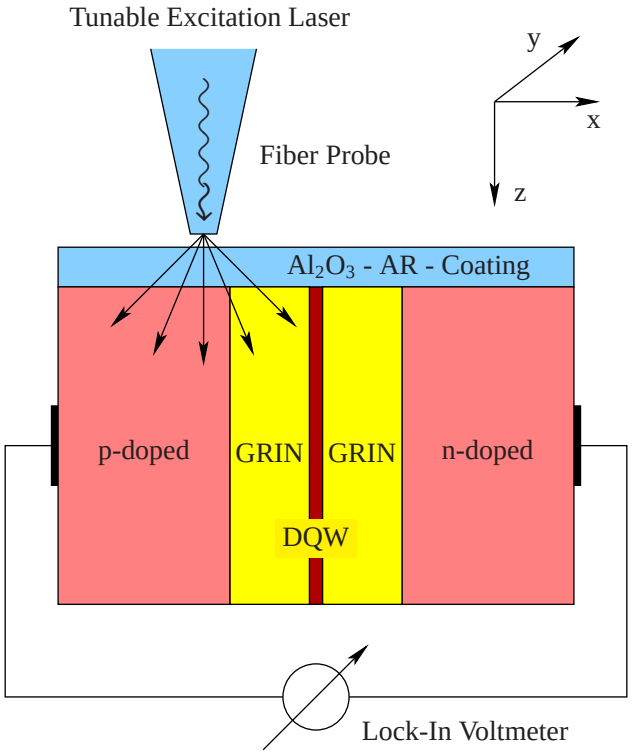

Figure 4.14: Scheme of the NPC experiment. During the scan the to sample distance was kept constant at 5±1 nm using an optical shear force setup described in subsection 3.2.2.

diode is excited by light transmitted through a 100-150 nm aperture at the end of a metal-coated NF fiber probe (107). The photoinduced current across the p-i-n junction is detected as a function of the tip position as the tip is scanned across the diode facet. The equipment used to perform the experiments described in secion 3.3.

Macroscopic FF PC spectra, recorded with a spatial resolution of about 500 μm, show a characteristic signal increase at an excitation energy of 1.53 eV that results from the onset of the interband absorption from the first heavy hole to the first electron subband in the DQW. The more than two orders of magnitude smaller contribution to the PC signal at energies below 1.53 eV is related to light absorption by defects inside the waveguide (108). The FF PC spectra are similar for all three diodes. The excellent agreement between FF and spatially integrated NPC spectra ensures that NF scans allow to analyze the microscopic origin of FF PC spectra (109).

Two-dimensional NPC images have been recorded for the three laser structures for different excitation wavelengths. A representative image for diode A and excitation at 1.53 eV, slightly above the onset of the DQW absorption, is shown in Fig. 4.15 (a). All images are rather homogeneous as the tip is scanned along the *y* axis, parallel to the DQW layer, while pronounced spatial variations in PC

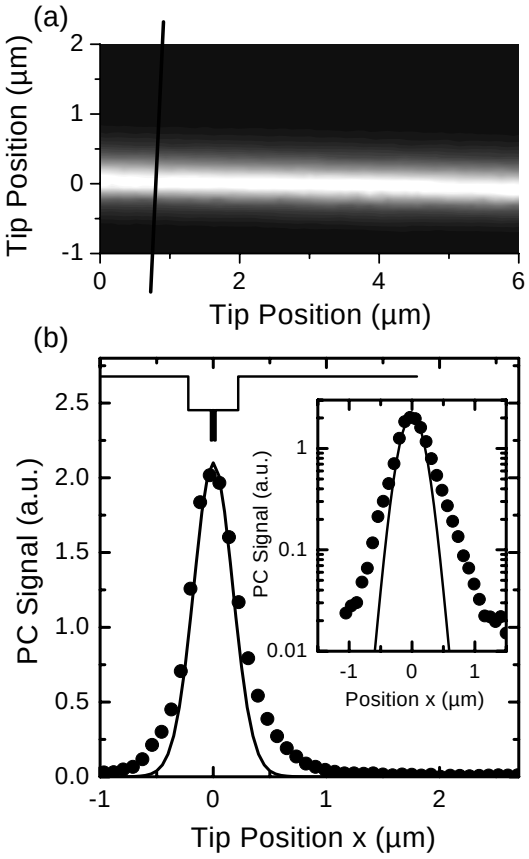

Figure 4.15: (a) Two-dimensional NPC image of a laser diode with a 0.44 μm wide SIN waveguide (diode A). (b) Cross section of the NPC signal along the *x* axis, perpendicular to the active layer (circles: experimental, solid line: simulation). Inset: NPC singal on a logarithmic intensity scale.

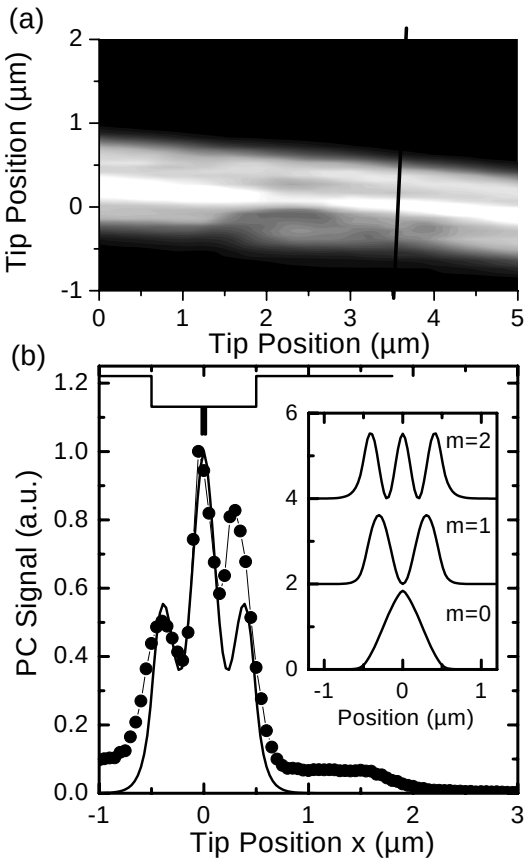

Figure 4.16: (a) Two-dimensional NPC image of a laser diode with a 1 μm wide SIN waveguide and a centered DQW (diode B). (b) NPC trace along the x axis, perpendicular to the active layer (circles: experimental, solid line: simulation). Inset: $|E_m(x)|^2$ for the three guided modes of the laser waveguide.

signal intensity are observed along x, perpendicular to the DQW layer. Figure 4.15 (b) shows a cross section through the two-dimensional image for diode A (440 nm wide SIN waveguide) along the x axis. In the inset the data are displayed on a logarithmic intensity scale. We observe a single narrow NPC peak with a full width at half maximum (FWHM) of 400 nm, which is centered at the DQW position. Outside the waveguide region, the signal intensity decays exponentially with increasing x [Fig. 4.15 (b), inset]. Different NPC images are observed for diode B (1 μm wide SIN waveguide, centered DQW) for the same excitation energy of 1.53 eV, Fig. 4.16. The cross section along x shows a narrow peak with a FWHM of about 250 nm centered at the DQW position. This peak is surrounded by two weaker, slightly broader and asymmetric side peaks that are separated by 400 nm from the central peak. The width of the entire triple peak structure is about 1 μm, corresponding to the width of the SIN waveguide layers. For diode C (1 μm wide SIN waveguide, off-centered DQW, Fig. 4.17), the triple peak is almost entirely washed out and the NPC signal consists of only a single broad peak with a width of 950 nm.

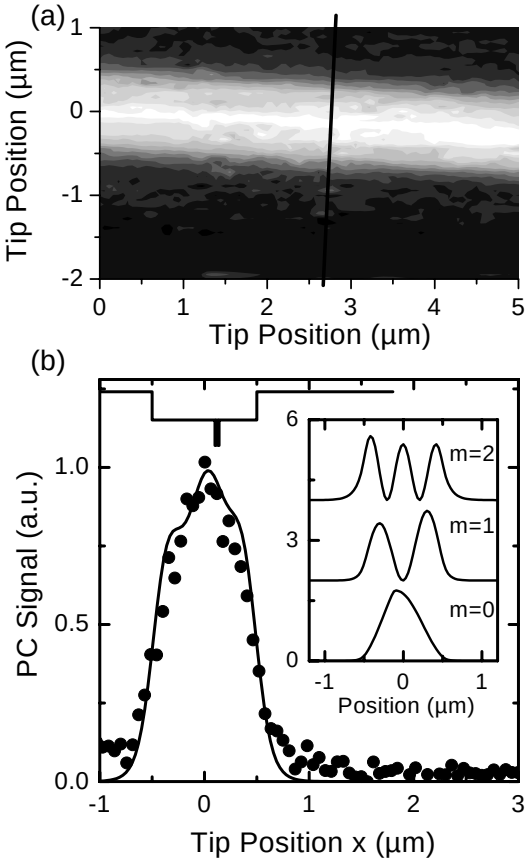

Figure 4.17: As Fig. 4.16 for diode C with a 1 μm wide SIN waveguide and a DQW that is off-centered by 120 nm.

4.3.4 Theoretical model of the NPC experiment

In these experiments, the light transmitted through the NF fiber probe constitutes an excitation source of subwavelength dimension that is coupled into the laser waveguide. The p-i-n junction of the laser diode then serves as a local photodetector, which detects the light intensity that is coupled into and absorbed within the waveguide. For metal coated NF aperture probes, the electromagnetic field distribution directly below the aperture consists of a superposition of both evanescent and propagating waves (110). The contribution from evanescent waves decays in the Al_2O_3 AR coating between NF probe and diode layers. Therefore, the electric field distribution $E_{in}(x, y)$ in the plane of the antireflection/diode interface, $z = 0$, that is coupled into the laser waveguide consists basically only of propagating fields with a maximum lateral component $k_{lat} = \sqrt{k_x^2 + k_y^2}$ of the wavevector $\mathbf{k} = (k_x, k_y, k_z)$ of $2\pi n_{\mathrm{AR}}/\lambda$. Calculations of the field propagation through the multilayer structure within a matrix transfer formalism show that the corresponding spatial intensity $I_{in}(x, y) = |E_{in}(x, y)|^2$ is well described by a Gaussian profile with a FWHM of 200 nm centered at the tip position. Thus, the effective spatial resolution in the experiment is not limited by the aperture diameter but is close to the diffraction-limited resolution defined by the refractive index of the AR coating, n_{AR}. This approximation seems valid if the AR layer thickness exceeds the decay length of the evanescent field $\lambda/2\pi n_{\mathrm{AR}}$. The efficiency for coupling $|\eta_m|^2$ into a mode m of the waveguide is then given by

$$|\eta_m|^2 = \frac{\left| \int\int_{-\infty}^{\infty} E_{in}(x, y)\, E_m^*(x, y)\, \mathrm{d}x\, \mathrm{d}y \right|^2}{\int\int_{-\infty}^{\infty} |E_{in}|^2\, \mathrm{d}x\, \mathrm{d}y \int\int_{-\infty}^{\infty} |E_m|^2\, \mathrm{d}x\, \mathrm{d}y}, \tag{4.45}$$

with $E_m(x, y)$ being the electric field distribution of waveguide mode m in the plane $z = 0$ (111). The coupling efficiency therefore maps the square of the overlap integral of the incident electric field E_in and the field profile of mode m. The field distribution inside the waveguide $E_w(\mathbf{r})$ can be decomposed into a superposition of a finite number of guided modes, $Eg_{m,i}(\mathbf{r})$, and a quasicontinuum of unguided or radiation modes, $Er_{n,i}(r)$:

$$Ew(\mathbf{r}) = \sum_{m,i} a_{m,i} Eg_{m,i}(\mathbf{r}) + \sum_{n,i} a_{n,i} Er_{n,i}(\mathbf{r}), \tag{4.46}$$

where $i = \mathrm{TE}, \mathrm{TM}$ denotes the mode polarization. For weakly guiding optical waveguides, the spatial mode profiles for TE and TM polarization are almost identical. We therefore restrict the following discussion to $i = \mathrm{TE}$ polarization. The mode profiles $Eg_m(\mathbf{r})$ and $Er_n(\mathbf{r})$ are obtained by solving Helmholtz equation. We find that the waveguide of diode A supports two guided modes. The intensity profile [proportional to $|Eg_m(\mathbf{r})|^2$] of the fundamental mode $m = 0$ is shown in the inset in Fig. 4.15. The increase in waveguide thickness from 440 nm to 1 μm in

diodes B and C increases the number of bound modes to three. Intensity mode profiles are shown in the insets of Figs. 4.16 and 4.17, respectively. Note that the off-center QW position in diode C breaks the lateral symmetry of the mode profiles. The spatial structures of these profiles have typical dimensions of 0.4 μm, larger than the width of I_{in}. Therefore, the efficiency of the coupling into the respective modes depends sensitively on the position of the NF-probe within the waveguide.

To simulate the NF images one has to evaluate the contribution of each excited waveguide mode to the PC signal. We assume that the PC signal is only generated by carrier absorption in the DQW layer. Around the laser photon energy, the FF PC spectra vary linearly with the DQW absorption coefficient (109). Thus absorption of propagating modes m is characterized by an effective absorption coefficient $\alpha_{eff,m} = -\mathrm{Im}(k_{z,m}^2)/[2\mathrm{Re}(k_{z,m})]$ which is defined by the overlap integral of mode profile and DQW and the DQW absorption coefficient. Due to the presence of the AR layer the contribution of evanescent modes is negligible. The total PC signal can then be written as the superposition of the contribution from guided and radiation modes as $I_{PC} \propto \sum_m |\eta_m|^2 \alpha_{eff,m} + \sum_n |\eta_n|^2 \alpha_{eff,n}$. The second term, the contribution from radiation modes, gives rise to a spatially slowly varying background signal for tip positions within the waveguide and cladding layers. Our simulations and experiments show that this contribution is significantly smaller than that from guided propagating waveguide modes[2]. The NPC images therefore map the space-dependent coupling of a localized light source into propagating guided modes of the laser waveguide and their absorption by the active DQW layer.

4.3.5 Experiment vs. theory notes

For active layers positioned in the center of the waveguide only modes of even symmetry are effectively absorbed, with a typical absorption length of about 10 μm. Therefore, for diode A only the ground mode $m = 0$ contributes to the NPC signal. The width and overall shape of the experimental NPC trace are well reproduced (Fig. 4.15). A closer inspection shows that the experimentally observed exponential decay of the signal for excitation outside the waveguide is weaker than the calculated one. This decay length is a sensitive measure of the difference in refractive indices between waveguide and cladding layers (98; 99). This difference is obviously overestimated in our simulations that did not consider the effect of doping on the refractive indices of the core/cladding layer.

As the number of guided modes is increased by changing the waveguide width to 1 μm, the NPC signal of diode B shows a superposition of $m = 0$ and $m = 2$ modes. Its overall shape is again well reproduced. Not reproduced is the pro-

[2]The contribution from radiation modes is expected to become significant if the thickness d_z of the waveguide structure (600 μm in our diodes) is much shorter than $1/\alpha_{eff,m}$ (~10 μm). In such thin waveguides, the propagating mode contribution reduces linearly with d_z and becomes less dominant as d_z approaches the wavelength of the light.

nounced asymmetry of the experimental trace. Such an asymmetry may only be obtained in our simulations by assuming slightly different refractive indices in the $p-$ and $n-$doped cladding layers. An off-center position of the active DQW layer (diode C) induces absorption of all three bound modes, $0 \pm m \pm 2$. Consequently, the pronounced triple peak structure observed for diode B is almost entirely washed out, as observed experimentally and theoretically (Fig. 4.17).

4.4 Nanostack®

4.4.1 Monolitic stacked lasers

Monolithic diode laser stacks have been first proposed several years ago by van der Ziel and Tsang (112) and are practically implemented by several groups (113; 114; 115; 116). In this kind of interband-cascade structures N p-n junctions are grown on top of each other. Thus a $(pn)^N$ structure is obtained. Applying an external voltage to this junction stack each second junction is forward biased, whereas the remaining junctions are reverse biased. The resulting unwanted current blocking effective for the whole device can be almost completely cancelled by designing the backward biased junctions as specially designed tunnel junctions. This is achieved by introducing extremely high p- and n- doping levels ($\sim 10^{19}$ cm^{-3}) close to the region where the conductivity type inverts. So in this region electrons and holes coexist at similar energy and are spatially separated only by an ultrathin depletion layer. This presents ideal conditions for carrier tunneling. It has been shown that the electro-optical behavior of such stacks scales is almost perfect; i.e., voltage drop and slope efficiency are multiplied by N whereas the threshold current remains almost unchanged (113; 114; 115; 116). This is maintained up to operation conditions where thermal effects become relevant. Thus such structures meet requirements of applications as high-brightness coherent radiation source for pulsed operation conditions. Additional technological advantages are related to the reduced operation currents that need to be provided by the power supply.

Generally, the design of multiple, lateral and vertical, structures such as arrays and stacks, are linked to reduced reliability figures. Different kinds of structural inhomogeneity are connected to these designs, e.g., non-equilibrium carrier concentration, light flux or temperature. Finally this results in an inhomogeneous thermal load. Therefore it is particularly important for monolithic laser stacks to appear either uniform or to compensate for inhomogeneities that are involved in the particular device architecture.

The main topic with which we will deal in this subsection is a uniformity analysis of monolithic stacked laser structures carried out by NSOM. We apply three different NSOM-based techniques, namely photoluminescence, photocurrent and analysis of the laser emission of the device itself. The results allow drawing conclusions regarding the homogeneity of the stacks within the devices.

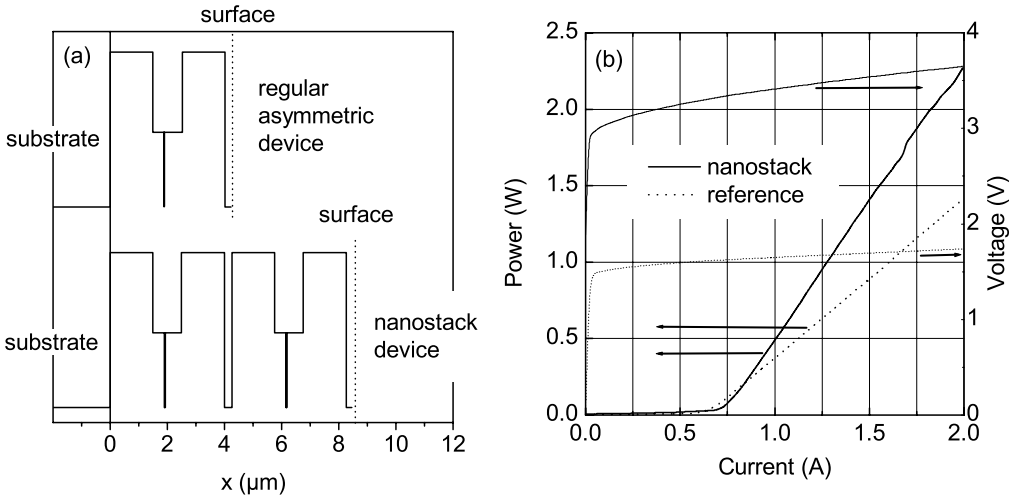

Figure 4.18: (a) Diagrams of the spatial variation of E_g along growth direction of the epilayers for a regular asymmetric LOC device (top) as well as for a Nanostack® structure. (b) L-I-V curves for a regular diode laser based on an asymmetric LOC structure and a Nanostack® device.

4.4.2 Stacked laser samples and experimental technique

The 808 nm (1.53 eV) emitting high-power laser diodes investigated are based on a standard asymmetric large optical cavity (LOC) structure grown by MOCVD method (see subsection 4.1.1). Devices involves a 1 μm thick $Al_{0.3}Ga_{0.7}As$ waveguide (energy gap, E_g = 1.8 eV) and an InAlGaAs DQW section located inside the waveguide but off-centered by 120 nm. The cladding material is $Al_{0.6}Ga_{0.4}As$ (E_g = 2.2 eV). A schematic diagram of the spatial variation of E_g is given in Fig. 4.18 (a) on top, cf. Ref. (117). Two such standard structures grown on top of each other and separated by a specially designed tunnel section form the Nanostack® devices shown in Fig. 4.18 (a) bottom. Fig. 4.18 (b) gives the L-I-V curves for a Nanostack® device together with data of a regular reference device. The data indicates that the stacked device scales nearly ideally, i.e., voltage drop and slope efficiency double whereas the threshold current is almost maintained. All investigated Nanostack® devices are from the same wafer and are packaged p-side down on standard *C*-mounts in order to secure optimized heat removal from the optically active regions.

The NSOM system used is based on a commercial Topometrix Aurora system. The setup is used in two different operation modes, namely in the luminescence collection and in the excitation mode. All experiments are performed at room temperature. Details of the experimental setup and the techniques employed are described in section 3.3 and in Ref. (118; 119; 120).

In the collection mode we detect either the electroluminescence or laser emission from the device, operated either below or above threshold, respectively, or the photoluminescence signal (118) excited by different excitation sources. We use the 442 nm line of a HeCd laser (E = 2.8 eV) and a tunable Ti:sapphire laser

that allows for excitation in the 1.48 – 1.75 eV range. These lasers are coupled into the fiber. The luminescence is collected trough the same fiber and detection is implemented by a liquid nitrogen cooled CCD-camera operating in single photon counting mode or an Si:avalanche photodiode. Thus contributions by diffused carrier pairs are reduced, e.g., compared to photocurrent experiments. This experimental approach ensures a spatial resolution of better than 150 nm. For details, see Ref. (120).

In excitation mode experiments the device itself serves as the detector for the induced PC (119; 107) which is fed into the NSOM electronics in the same way as the detector signals in the PL experiments. Under these conditions, compared to the PL experiments, the spatial resolution is poorer. Nevertheless, as we demonstrate in section 4.3 (see also (119)), PC structures in diode laser waveguides, separated by 400 nm are clearly resolved as single peaks.

4.4.3 Electroluminescence and laser emission experiment

Fig. 4.19 shows electroluminescence (a) and laser emission maps (c) obtained from a Nanostack® device. For the electroluminescence measurement the device was operated with 40 mA pulses having a duration $\tau = 410\ \mu$s at a repetition rate of $f = 1.22$ kHz resulting in a duty cycle of 50%. For the laser emission experiment at 1 A a duty cycle of 0.0024% was chosen ($f = 1.22$ kHz, $\tau = 20$ ns). Thus despite the increased operation current the integrated thermal load for the laser experiment is reduced by a factor of 820. From such laterally homogeneous maps we extract linescans that are shown in Fig. 4.19 (b) and (d) for electroluminescence and lasing, respectively. Obviously there is an asymmetry between the two sections that is even stronger pronounced for the laser emission experiment. Thermal effects do not account for this as indicated by Figs. 4.19 taking into account the substantially reduced thermal load in the laser experiment. Extra measurements at different excitation currents confirm this, too.

4.4.4 Photoluminescence experiment

In order to get more details on the observed asymmetric luminescence behavior we change the luminescence excitation mechanism by switching from electrical to optical excitation, i.e., to a PL experiment. Fig. 4.20 shows PL data for two different excitation energies, namely 2.8 eV (a) and 1.69 eV (b). For both experiments the spectral detection window is set to the QW emission energy at 1.53 eV. Excitation with 2.8 eV implements surface excitation for *all* regions of the structure (including the cladding layer) [cf Fig. 4.18 (b)], whereas the 1.69 eV photons excite exclusively the QWs and the GaAs ($E_g = 1.424$ eV) specially designed tunnel junction region.

Obviously the asymmetric behavior observed in the device emission is also present for external photoexcitation regardless of the QW is directly excited [cf.

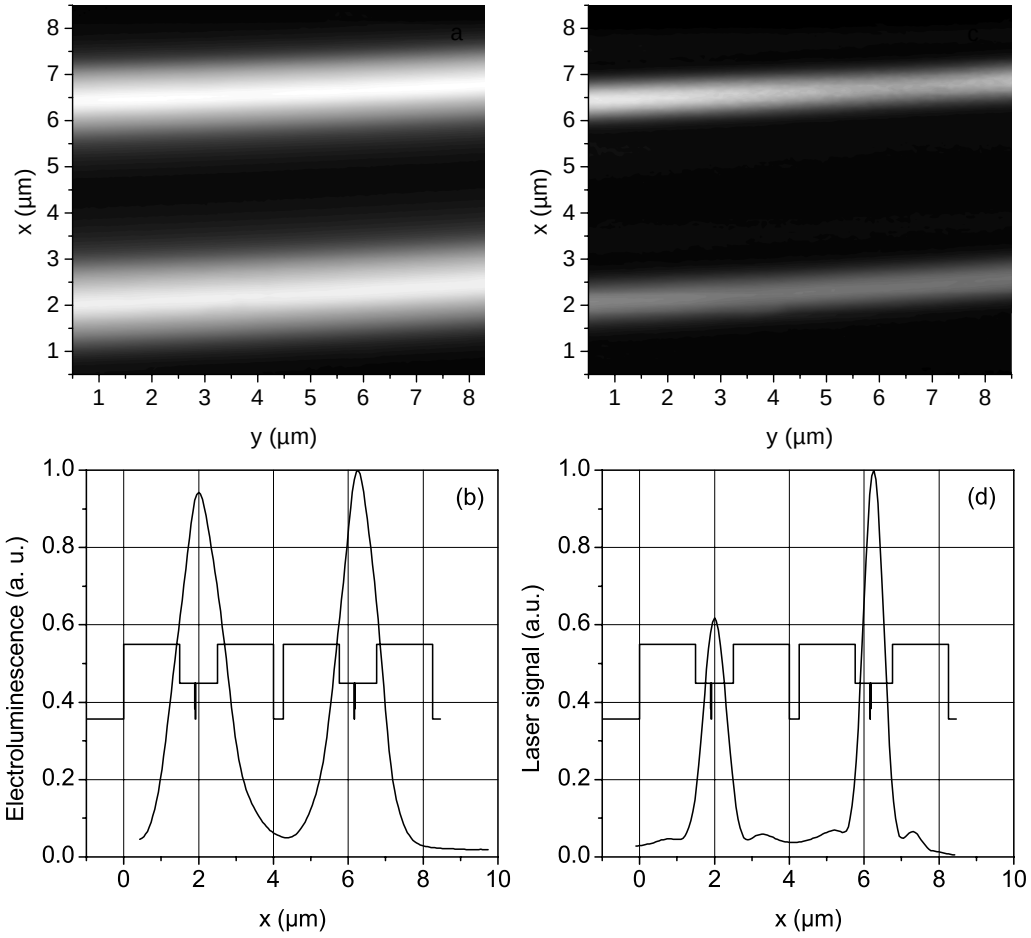

Figure 4.19: Electroluminescence (a) and laser emission (c) maps obtained from a Nanostack® device. From such maps we extract linescans that are shown in (b) and (d). The pulsed device operation conditions are for electroluminesce $f = 1.22$ kHz, $\tau = 410\,\mu$s, and for laser emission $f = 1.22$ kHz, $\tau = 20$ ns. Thus the integral thermal load is by a factor of 820 lower for the lasing experiment compared to the electroluminescence one.

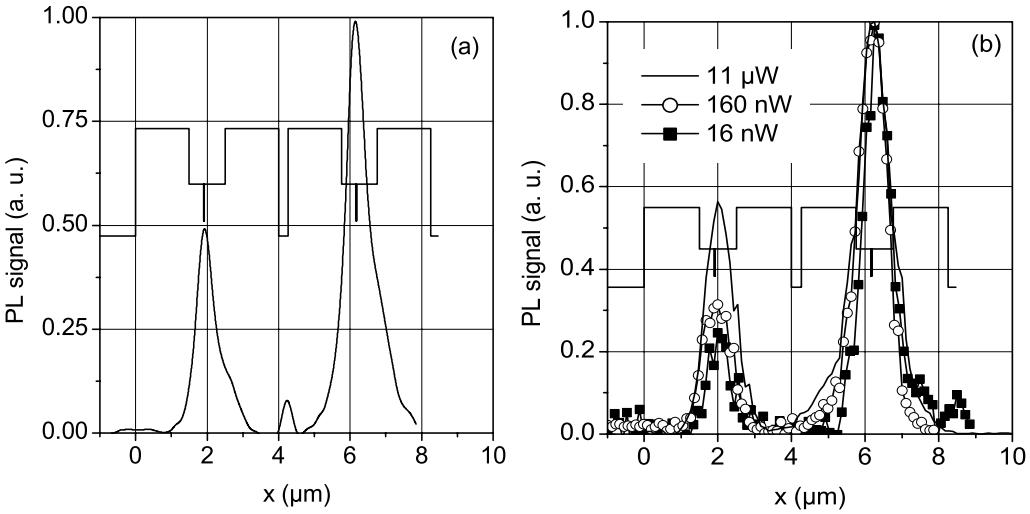

Figure 4.20: PL line scans across the Nanostack® laser structure: (a) The excitation energy is 2.8 eV causing surface excitation. Thus the information depth in this experiment is exclusively determined by diffusion. The detection window is set to the QW emission photon energy of 1.53 eV. (b) The excitation photons of 1.69 eV energy selectively excite the QW region and become absorbed within about 100 μm (information depth). The detection window is set to the QW emission photon energy of 1.53 eV. The parameter is the excitation power.

Fig. 4.20 (b)], or predominantly indirectly populated by carriers generated originally within the waveguide [cf. Fig. 4.20 (a)]. Additionally Fig. 4.20 (b) shows an excitation intensity dependence of the PL emission for exclusive QW excitation. For very low excitation densities the observed asymmetry becomes further enhanced.

4.4.5 Photocurrent experiment

Now we present the PC experiments. Data are shown in Fig. 4.21 and again indicate an asymmetric behavior of the two segments of the stack. For these experiments, however, the stronger response comes from the segment close to the substrate, i.e., the one that shows the poorer luminescence behavior. Furthermore, the asymmetry is less pronounced for surface excitation [2.8 eV, cf. Fig. 4.21 (a)] compared to selective QW excitation [1.59 eV, cf. Fig. 4.21 (b)].

The specific shape of the scan excited at 780 nm (1.59 eV) is explained by the mode structure of this particular LOC waveguide design as it was quantitatively analyzed in sections 4.2 and 4.3 for similar waveguides being incorporated into devices with only one waveguide (119).

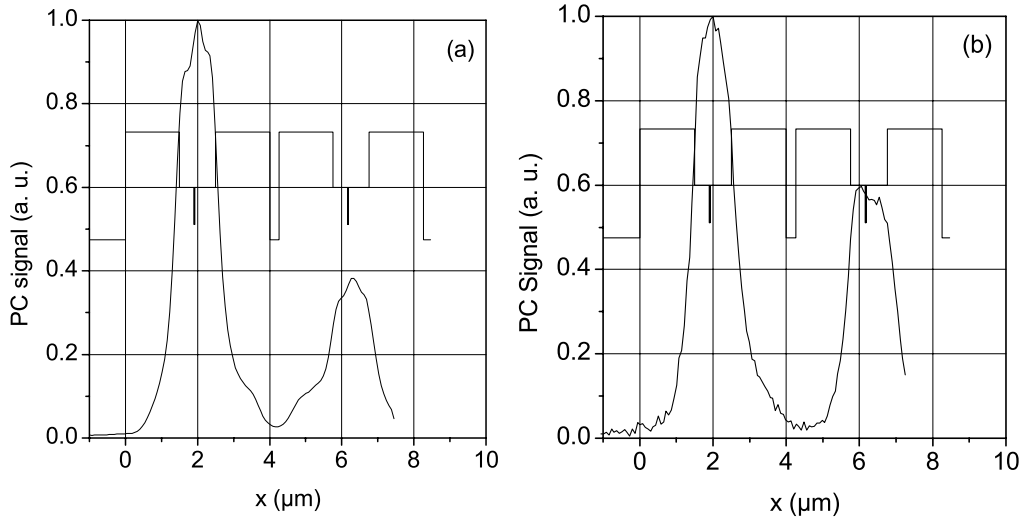

Figure 4.21: PC line scans across the Nanostack® laser structure. (a) The excitation resonant to the QW energy is 1.59 eV. The multiple structures are due to the excitation of the three confined waveguide modes (see Ref. (120)). (b) The excitation resonant to the QW energy is 2.8 eV causing surface excitation.

4.4.6 Signal generation mechanisms

In order to consistently interpret the data obtained with the different techniques, we summarize in brief the relevant signal generation mechanisms.

(i) The room temperature QW PL signal is proportional to $\delta n \sim \tau$. Here δn is the non-equilibrium concentration of electron-hole pairs and τ their total lifetime. For very low excitation densities carrier trapping into a finite number of defects might be effective. This results in a reduction of PL intensity at low excitation levels. Moreover, saturation of the trapping at higher excitation densities might cause a superlinear dependence of the PL intensity on δn.

(ii) The PC signal is proportional to $\delta n \times \mathrm{grad}(V)$, where V is the effective potential experienced by the carrier pairs. V is partly determined by the band gap variation (Fig. 4.18(a)), but–at the very high doping levels in this devices–doping has an even stronger influence. So large grad(V) values are expected in the depletion layers of the waveguides. The effect of the nominally even larger grad(V) value at the specially designed tunnel junction, however, is likely to be compensated by the tunnel effect itself which results in a suppression of spatial carrier separation and hence in a reduction of the PC contribution from the specially designed tunnel junction. If this would not be the case the junction would not operate properly. Fig. 4.18, however, clearly demonstrates the excellent performance of this crucial device part.

(iii) The electroluminescence as well as the laser signal depend in a very complex way on several parameters, among them also δn and grad(V). Finally, this complexity is the reason why we are not able to restrict our study to the emission properties of both device sections, (cf. Figs. 4.19). We note that the potential V

in the emission experiments differs substantially from the one relevant in (i) and (ii), since the external voltage levels the contribution that originally arises from the doping. The voltage drop across the device (Fig. 4.18(b)) is a measure of this potential leveling.

4.4.7 Explanation of the laser stack asymmetric behavior

First we discuss effects that might be inherent for the specific NSOM experiments. One may argue that the different behaviors of the two segments in Figs. 4.19, 4.20, 4.21 arises from the complex twin-waveguide architecture including coupling effects between both laser segments, or reflection at the p-contact metallization. The different behavior of PL and PC ratios, where at least the generation process of carrier pairs is similar, is in clear contrast to this interpretation. Second, the PL and PC experiments with high-energy surface excitation (cf. Figs. 4.20 and 4.21) show almost the same intensity ratio for the signals from both sections as those experiments for resonant QW excitation, where the three confined waveguide modes (for details see section 4.3 and (119)) are resolved as clear structures in the spatial PC scans. Thus interference of waveguiding effects on our data can be ruled out, too.

Electroluminescence as well as laser emission show weaker emission intensity from the QW in the device section that is situated closer to the substrate. This could be explained by an increased defect concentration within or in the vicinity of these QWs. The increase of asymmetry in the lasing experiment compared to the electroluminescence is then explained by the nonlinearity of the laser L-I curve.

The PL experiment with exclusive excitation of the QWs within the "bulk" of the waveguide (information depth ~100 μm) [cf. Fig. 4.20 (b)] confirms the assumption of an increased defect concentration within the device section that is situated closer to the substrate by proving that this effect is independent on the generation mechanism. Furthermore, the intensity dependence of the PL indicates a behavior that is indicative for trap saturation within the device section that is situated closer to the substrate.

Independently, the PC experiment provides additional confirmation: If one compares Fig. 4.21 (a) and (b) one finds that the asymmetry is less pronounced for surface excitation (2.8 eV). Since grad(V) is almost independent on the depth where δn is created (defined by the excitation energy) this is an indication that the enhanced nonradiative recombination appears rather in the "bulk" of the waveguide than at the surface of the structure. This argument is reasonable since it appears justified to assume that the facet status of both laser segments is similar.

Despite of the different signal ratios of both device sections obtained in PL and PC, both sets of data clearly indicate an enhanced recombination efficiency in the QW in the device section next to the substrate. The reduced PL intensity reflects the trapping of carriers into defect states. PC and intensity dependent PL data identify the bulk of the waveguide as the main location of the defects. Thus

the assumption of an increased trap concentration in one of the laser segments consistently explains the asymmetric behavior of the laser stack.

4.4.8 Nanostack® summary

We investigate the optoelectronic properties of monolithically stacked high power diode lasers including two nominally identical waveguides separated by a specially designed tunnel junction. NSOM analysis straight and separately addresses the properties of both segments. Device emission, namely electroluminescence and lasing, as well as PL and PC data are recorded and consistently discussed. It is shown that the coupling between both waveguides is marginal.

We find slightly reduced laser emission from the laser segment that is situated closer to the substrate. We show that this is not caused by thermal effects or by effects caused by the complex architecture of the two waveguides but most likely due to a larger trap concentration within or in the vicinity of the QWs of this laser waveguide. Furthermore, we show that in the unbiased devices the potential gradient in this segment is significantly larger than in the one close to the heat sink. The latter effect is not necessarily directly connected to the detected different trap concentration, however, both might be linked to the doping.

4.5 Novel technique for determination of surface recombination velocity and diffusion length

4.5.1 Problem particularities and experemental details

Surface recombination in semiconductors describes the annihilation of charge excitations near the surface of a crystal. Phenomenologically, SR is parametrized by a surface recombination velocity that is an important factor in characterizing the surface properties. Microscopically, the main determining factor for SRV is the recombination center density at the surface, which is related to the overall bulk quality of a sample (121). In modern optoelectronic devices, SR is playing a more and more critical role as devices become smaller. In high power diode lasers, e.g., it is well known that SR increases the thermal load on the mirror surfaces, giving rise to thermal runaway and eventually limiting the lifetime of the device (122). With increasing miniaturization also the diffusion length (DL) is becoming a critical factor (123), defining the region of the device that is affected by SR.

A variety of different techniques have been used to determine the SRV and DL (124; 125; 126; 127; 128) but most of these measurements suffer from a lack of adequate spatial resolution. To unambiguously determine DL, it is essential to probe non-equilibrium carrier concentrations with a resolution that is less than DL, i.e., typically on the order of 1 μm in direct band gap III/V semiconductors. The same holds true for SRV, because one needs to differentiate between the surface region and the bulk region as defined by the characteristic length scale of the

DL.

PL based techniques are powerful and nondestructive probes of non-equilibrium carrier distributions and their relaxation dynamics (129). Yet, the spatial resolution of FF PL is often larger than DL. NSOM, providing spatial resolution on the order of 100 nm, overcomes these limitations and thus is a particularly powerful tool for the nondestructive analysis of SR and diffusion processes in semiconductor devices.

We have performed a *simultaneous* nanoscopic measurement of SRV and DL on a semiconductor QW. NF microscopy is used to probe PL in the vicinity of a cleaved and plasma-treated edge of a QW sample. By comparison to a two dimensional diffusion model, both SRV and DL are extracted as a function of excitation intensity.

The sample is a single InGaAs QW diode laser structure grown by MBE on a Si-doped GaAs substrate. The 7 nm $In_{0.16}Ga_{0.84}As$ QW is buried 220 nm below the sample surface. It is clad from both sides between 10 nm GaAs layers, surrounded by 200 nm $Al_{0.2}Ga_{0.8}As$ layers. After growth, the sample was cleaved in air in the direction perpendicular to the QW plane. The cleaved surface was exposed to a plasma cleaning and coating process in order to improve the surface quality (125; 130).

Spatially resolved NF PL experiments are performed at room temperature [Fig. 4.22]. The microscope described in section 3.3 is used in the illumination/collection geometry: the He-Ne excitation laser (photon energy 1.96 eV) is transmitted through the NF fiber probe and PL emitted from the sample is collected through the same fiber. An interference filter separates QW PL emission [centered at 1.30 eV, 15 meV FWHM] from the excitation laser. Uncoated, chemically etched NF fiber probes are used, providing 150 nm spatial resolution and transmission efficiencies close to unity (131; 132). QW luminescence lifetimes are recorded with a synchroscan streak camera using a sub-100 fs Ti:sapphire laser centered at 1.57 eV as the impulsive excitation source.

Two dimensional NF PL images $PL(x,y)$ are recorded for different excitation powers between 3 and 300 μW coupled into the NF fiber probe. Images at 10 and 300 μW are shown in Fig. 4.22 (b) and (c), respectively. Strong QW PL is observed at distances $x > 1$ μm from the surface ($x = 0$). The PL decays gradually as the tip approaches the surface and this gradual decrease is homogeneous along the y-axis, parallel to the sample edge. It is clear that the high intensity PL [Fig. 4.22 (c)] is spatially more uniform than the low intensity PL. In order to highlight the difference, we provide line-leveled images $\Delta PL(x,y)$ of each Figure in Fig. 4.22 (d) and Fig. 4.22 (e), respectively. We subtract the average value of each column and then replot the image, $\Delta PL(x,y) = PL(x,y) - \langle PL\rangle(x)$, where $\langle PL\rangle(x)$ denotes the signal average along the y-axis. This shows strong local fluctuations in the low power PL image [Fig. 4.22 (d)], in contrast to the almost flat high power image [Fig. 4.22 (e)]. Such differences can be explained by the saturation of trap states at high excitation levels, as discussed below.

As the PL images are homogeneous along y, we compare in Fig. 4.23 only

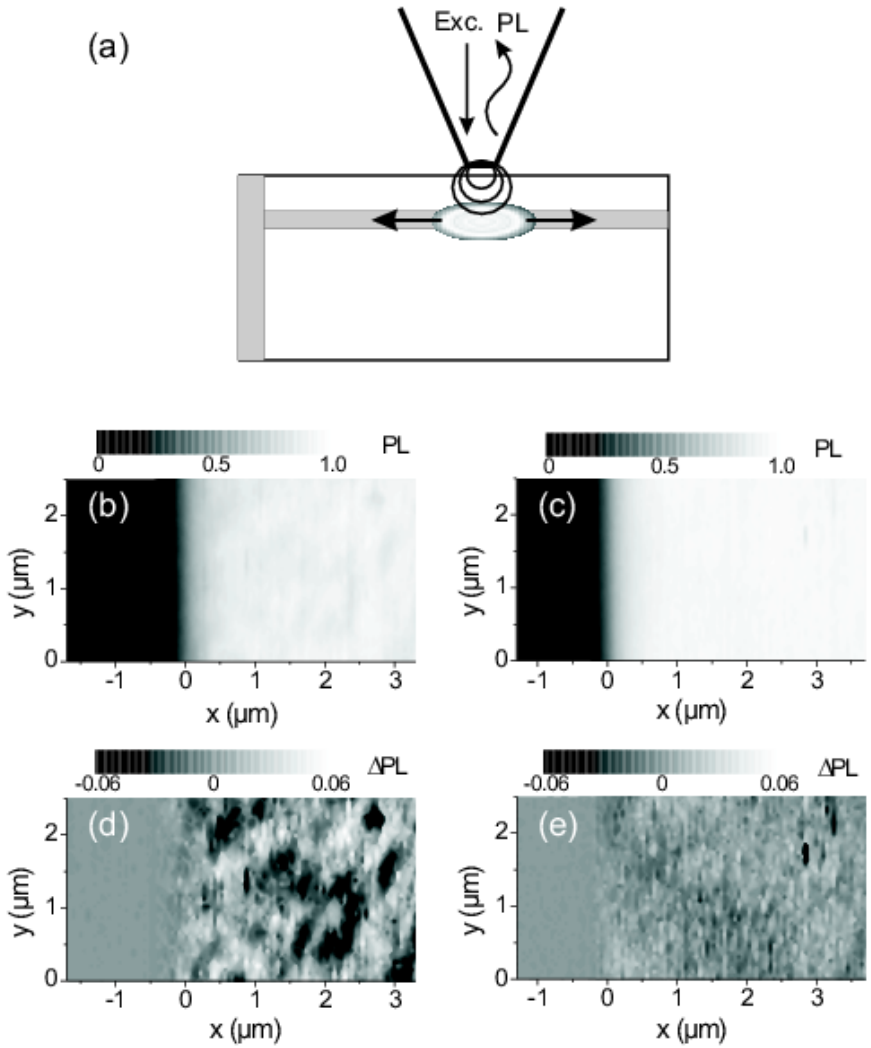

Figure 4.22: (a). Schematic of the NF PL experiment in illumination/collection geometry. The excitation laser (photon energy 1.96 eV) is transmitted through a chemically etched, uncoated NF fiber probe. QW PL is collected through the same fiber and detected by a single photon counting system. (b) Two-dimensional, NF QW PL intensity $PL(x,y)$ near the edge ($x = 0$) of the QW sample. Excitation power $P = 10\ \mu$W. $PL(x,y)$ is color-coded and normalized unity. (c) As (b), $P = 300\ \mu$W. (d) Line-leveled image $\Delta PL(x,y)$ of (b), $P = 10\ \mu$W. (e) $\Delta PL(x,y)$ of (c), $P = 300\ \mu$W.

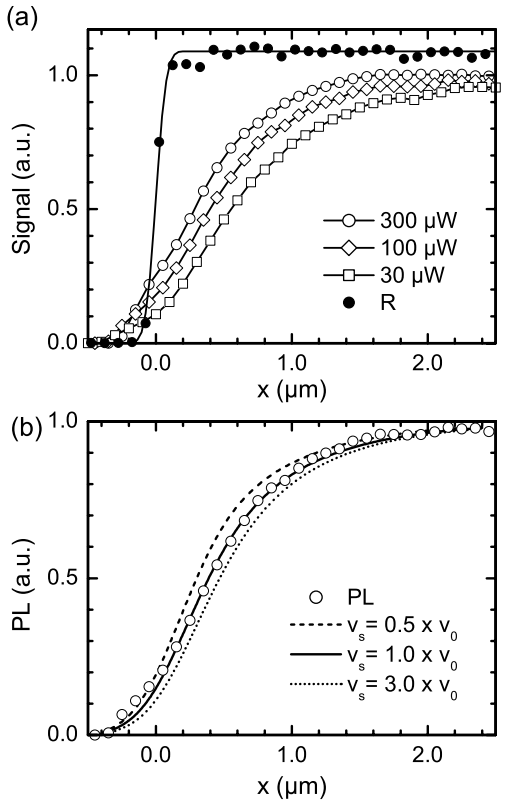

Figure 4.23: (a) QW PL line scans signal recorded at different excitation densities near the sample boundary: $P = 30$ μW (squares), 100 μW(diamonds), and 300 μW (open circles). NF reflectivity scan (full circles) showing the spatial resolution of the experiment. (b) QW PL line scan for $P = 100$ μW (circles), and simulations to the two-dimensional diffusion model [Eq. (1)]. Solid line: $v_s = v_0 = 1 \times 10^6$ cm/s, $D = 15$ cm^2/s, $\tau_{rec} = 1.7$ ns. Dashed and dotted lines are for $v_s = 5{\times}10^5$ cm/s and $v_s = 3 \times 10^6$ cm/s, respectively, and the same D.

PL line scans along the x axis near the edge at different excitation powers. A reflectivity measurement is included in the same picture (filled circles) in order to show the spatial resolution of our experiment. From this curve we can deduce a spatial resolution of 140 nm in the reflectivity measurement. In all PL profiles, the PL rises much more slowly as a function of x than the reflectivity, with a characteristic length scale of about 1 μm, approximately the DL of our system. Neglecting the finite spatial resolution, the PL intensity around $x = 0$ would be zero in the limit of an infinite SRV (perfectly absorbing boundary conditions). This PL intensity should increase as SRV, making the intensity close to $x = 0$ a sensitive measure of SRV.

4.5.2 2D-diffusion model

In order to quantitatively describe the intensity profile and to deduce SRV and DL, we use a standard steady-state diffusion model with a boundary at $x = 0$ where the SR acts as a sink for the carriers. For excitation at 1.96 eV, electron-hole pairs are generated in the $Al_{0.2}Ga_{0.8}As$ regions and are rapidly trapped into the QW. This trapping occurs on a time scale of few tens of ps, much faster than the QW luminescence life time τ_{rec} of ~1 ns, as confirmed by time-resolved PL. The photoexcited carriers trapped into the QW undergo diffusion within the QW plane. At room temperature, this diffusion is well described by an *ambipolar* diffusion model (133), with a characteristic DL $L_D = \sqrt{D\tau_{rec}}$, given by the ambipolar diffu-

sion coefficient D and τ_{rec}. We therefore assume identical electron and hole QW densities $n_e^{QW} = n_h^{QW} = n$, their spatial variation being described by the following two-dimensional diffusion equation:

$$D\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)n - \frac{n}{\tau_{rec}} + n_a\, g(x_0, y_0) = 0. \tag{4.47}$$

Here, $g(x_0, y_0) = (1/\sqrt{2\pi}\sigma)\mathrm{e}^{-[(x-x_0)^2+(y-y_0)^2]/2\sigma^2}$ is a local source term, describing carrier generation by the NF-tip positioned at (x_0, y_0). It is modeled by a Gaussian spatial profile (134) with a FWHM $\sqrt{2\ln(2)}\sigma$ of 300 nm. Due to finite depth of the buried QW layer, its width is larger than that deduced from the reflectivity measurement. The generated carrier density is given by the constant n_a. We numerically solve Eq. 4.47 with the following boundary condition at the sink, $x = 0$:

$$D\frac{\partial n}{\partial x}\bigg|_{x=0} = v_s \cdot n.$$

The locally detected PL intensity $PL(x_0, y_0) \propto \iint \mathrm{d}x\mathrm{d}y \cdot g(x_0, y_0) \cdot n$. From the spatial variation of the PL signal alone, only two parameters, the DL L_D and the product $(v_s \cdot \tau_{rec})$ can be deduced. To independently extract all three parameters of the model, i.e. D, τ_{rec} and v_s, additional information is needed. Therefore, QW luminescence lifetimes τ_{rec} are measured by time resolving the QW PL after fs-excitation at 1.57 eV [Fig. 4.24]. Lifetimes of about 2 ns, typical for this

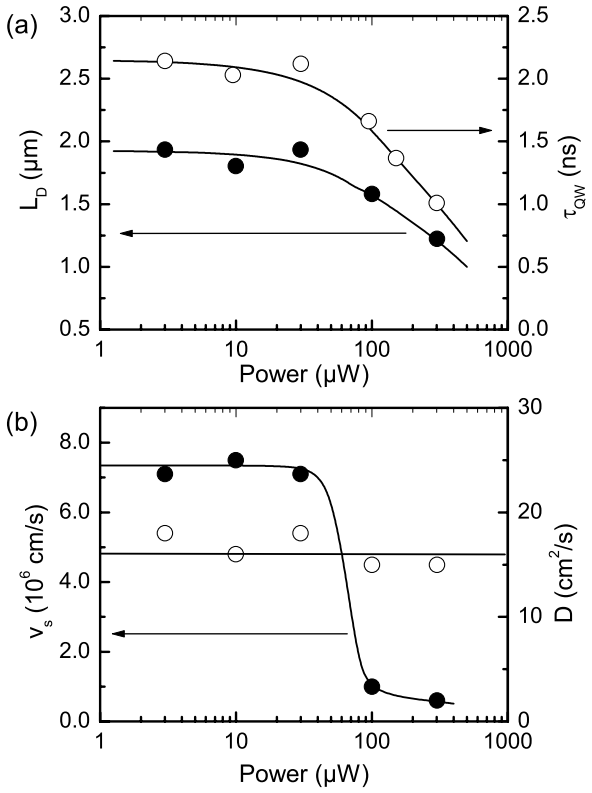

Figure 4.24: (a) DL L_D (open circles) deduced from the simulation of the NF PL line scans to the diffusion model and QW PL lifetime τ_{rec} measured by time resolved PL spectroscopy. (b) SRV (v_s, filled circles) and diffusion coefficient (D, open circles), extracted from the simulation of the PL line scans using the measured values of τ_{rec}.

kind of QW for diode laser applications (135), are measured at low excitation

densities. τ_{rec} decreases to about 1 ns toward the high-density regime. This trend corresponds to earlier experimental results on such QW (135; 136; 137).

In order to be able to fit experimental data we write a FORTRAN program that simulate PL NF experiment in excitation/collection mode. We locate the tip at the sample boundary and move it with the step of the mesh along the x-axis (perpendicular to the sample surface) (y-coordinate is fixed) inside the sample. At each point we calculate the steady-state carrier distribution by plug-in of a Gaussian initial distribution created by NF-tip (134) into a diffusion equation solver. Then, by convolution of calculated carrier distribution with the tip profile and assuming tip collection efficiency as 10% we obtain an integral value that correspond to the detected PL-signal.

The diffusion equation solver based on the use of the standard routine "sepeli" (138) that is the part of the "Fishpack" package of FORTRAN subprograms for the solution of separable elliptic partial differential equations (139). The subroutine "sepeli" solves for either the second-order finite difference approximation or a fourth-order approximation to a separable elliptic equation on a rectangle (140; 141).

$$F_a(x)\frac{\partial^2 U}{\partial x^2} + F_b(x)\frac{\partial U}{\partial x} + F_c(x)U + F_d(y)\frac{\partial^2 U}{\partial y^2} + F_e(y)\frac{\partial U}{\partial y} + F_f(y)U = G(x,y)$$

Any combination of periodic or mixed boundary conditions is allowed.

Before calling "sepeli" to solve the steady-state diffusion equation we need to adjust all coefficients in the equation as close to "1" as possible. This insures that underflow/overflow problems do not affect results. We implement this by switching to a dimensionless coordinates through substitution: $x_{dimless} = x/L_d$, $y_{dimless} = y/L_d$, $N(x_{dimless}, y_{dimless}) = n(x,y)/cn_0$, where L_d is a characteristic DL and cn_0 is a characteristic carrier concentration. In the normalized equation the dimensionless coefficients before the second derivative and the loss term are equal to "1".

4.5.3 Discussion

The cross-sectional PL intensities near the edge are then compared to the simulated PL profiles using the diffusion model [Fig. 4.23 (b)] with L_D and $v_s\tau_{rec}$ as independent parameters. The fitting is sensitive to both L_D and $v_s\tau_{rec}$. L_D is mostly determined by the larger x region, whereas $v_s\tau_{rec}$ affects the relative intensity at $x = 0$. The sensitivity of the simulation to v_s is illustrated. Over the entire density range, the general agreement between the experiment and simulation is quite satisfactory allowing deducing both fitting parameters. Similar to the lifetime, the DL is constant (2 μm) at low powers and decreases at powers of more than 30 μW. This decrease is clearly linked to the decrease in carrier lifetime. Within the experimental accuracy, the diffusion coefficient $D \approx 16$ cm^2/s is density independent [Fig. 4.24 (b)]. Using an Einstein relation, $\mu_{eff} = eD/kT$ (e is electron charge, kT is thermal energy) one derives $\mu_{eff} = 620$ cm^2/Vs at $T = 300$ K. This

value is close to the mobility of holes in $In_{0.16}Ga_{0.84}As$ (137), whereas electron mobilities are much higher. We conclude that hole diffusion determines the experimentally measured DL, and hence validating the assumption of an ambipolar diffusion model.

Now we discuss the values of v_s deduced from the simulation of the NF experiments [Fig. 4.24 (b)]. At low excitation powers $\leq$30 μW, the PL intensity around $x = 0$ is close to zero. In this regime, the sample surface can in good approximation be considered as a perfectly absorbing boundary. The value of $v_s = 7 \times 10^6$ cm/s given in Fig. 4.24 (b) is the minimum value of v_s that allows to satisfactorily fit the low PL intensity around $x = 0$. Clearly, our experiment is not sensitive to higher values of SRV. The experimental PL intensity at $x = 0$ increases substantially for the higher powers and hence the perfectly absorbing boundary description becomes inadequate. Values of v_s of 1×10^6 cm/s and 6×10^5 cm/s are deduced at powers of 100 and 300 μW, respectively (137). This density dependence shows that the saturation of nonradiative recombination near the QW surface is the microscopic mechanism behind the decrease of v_s at high densities. A microscopic assignment of this mechanism is beyond the scope of this thesis but it is likely that trap-like zero-dimensional centers are responsible (142). Their areal density can roughly be estimated from Fig. 4.24 (b) as 5×10^{11} cm^{-2}. It is interesting that this excitation density coincides with the disappearance of the lateral intensity fluctuations in the two-dimensional NF PL images (Fig. 4.22), i.e., at positions far away from the edge. Here, obviously, the microscopic mechanism causing these fluctuations is different. As τ_{rec} varies only slightly, radiative defects arising, e.g., from interface roughness (143), are responsible for the spatial fluctuations of the PL intensity.

In conclusion, in this section we describe a novel technique for the simultaneous nanoscopic measurement of SRV and diffusion constant in semiconductor nanostructures. Spatially resolved NF PL scans at the edge of a QW sample are compared to a two dimensional diffusion model and SRV and DL are deduced for various excitation powers. A pronounced decrease of SRV with increasing excitation power is a strong indication for saturation of nonradiative center near the sample surface.

Chapter 5

Plasmon nano-optics

This chapter is devoted to plasmon nano-optics. The objects under investigation are two-dimensional plasmonic crystals. Which are the optically thick metal film deposited on a dielectric substrate and perforated with a periodic array of nano-holes. At the beginning we will look at the properties of metalls that are responsible for the existence of plasmon excitations. Then we consider properties of surface plasmons (SP) on a flat surface and their dispersion relation. We will compare quasi-two-dimensional photonic crystals and two-dimensional plasmonic crystals. Fundamental problems in the first type of systems will be shown and solutions with the help of lasts. The numerical methods frequently used for the investigation of electromagnetic systems will be considered. Difficulties and benefits of time-domain and frequency-domain calculations will be shown.

The microscopic emission patterns of light transmitted through plasmonic crystals will be presented. We will show that the dominant contribution to the emission can be from the metal surfaces outside the holes. The NF pattern also reflects the symmetry of the SP waves that are polarization controlled. We will present the measurements that reveal that the emission from the metal surfaces can reach the FF region. We will show that the NF patterns are determined by the coherent interference of many SP waves, while efficient wavevector selection makes the FF pattern dominated by only the first two diffraction orders.

We will present the direct measurement of the damping of SP excitations in periodic nano-hole arrays. By probing coherent spatial SP propagation lengths of a few μm and an ultrafast decay of the SP polarization on a 10 fs timescale, we will demonstrate that the SP transmission peaks are homogeneously broadened by the SP radiative lifetime. We will show that the pronounced wavelength and hole size dependence of the damping rate points out that the microscopic origin of the conversion of SP into light is a Rayleigh-like scattering by the periodic hole array.

We will show that near-field patterns are governed by polarization and arise from a linear superposition of surface plasmon modes of all orders. Moreover, plasmonic crystalls, when excited near the surface plasmon resonance, can act as a vertical wave-guide to generate a light wave with most of its momentum perpendicular to its propagation direction. We will show that the efficient wave vector

selection generates spatial by sinusoidal waves that result from the interference between the zeroth and the first order wave vector components. When the wavelength of exciting light is below the lattice constant, near-perfect sinusoidal pattern can persists well into the far-field, up to 15 microns away from the surface.

5.1 Metallic particularities

5.1.1 What is different from dielectric

The main thing that is responsible that the optical properties of metal are so different from dielectrics is the complex refractive index of metal $n = \sqrt{\varepsilon + 2i\sigma/\nu}$. The large value of the conductivity σ mostly results in a large negative value of the dielectric constant. The oscillating electric field acting on free metallic electrons gives rise to a current 90° out of phase, representing a negative dielectric polarization. The resistivity of the metal results in the generation of heat through a small component of the current in phase, proportional to σ. Therefore one can consider the negative coefficient ε as the important characteristic and σ as a correction representing energy losses.

A negative value of ε means that the material is optically less dense than a vacuum. Then every wave striking such a medium will experience total reflection, very much as in the case of total internal reflection in transparent media. σ give rise to a partial absorption of that radiation, which penetrates the metal during the reflection process. Such energy losses can be completely ignored in systems that interact with electromagnetic field on a femtosecond time scale shorter than the damping time. The plasmon crystalls belongs to such systems and in their theoretical consideration conductivity/resistivity of the metal can be completely ignored.

5.1.2 Volume and surface plasmons

Many of the fundamental electronic properties of the solid state can be successfully described by single electrons moving in the periodic grating of atoms. Another approach is based on a plasma concept that assumes that free electrons of a metal are treated as an electron liquid of high density of about 10^{23} cm^{-3}, ignoring the lattice in a first approximation. From this approach, it follows that longitudinal density fluctuations, plasma oscillations, will propagate through the volume of the metal. Such "volume plasmons" have an energy $\hbar\omega_p = \hbar\sqrt{4\pi ne^2/m_0}$, where n is the electron density, of the order of 10 eV. Experementally, these volume plasmons are often exited by electron beams impinging on the metal.

An important extension of the plasmon physics is the concept of "surface plasmons" (SP). Maxwell theory shows that electromagnetic surface waves can propagate along a metallic surface or on metallic films with a broad spectrum of eigenfrequencies from $\omega = 0$ up to $\omega = \omega_p/\sqrt{2}$ depending on the wavevector $\mathbf{k}$. Their

dispersion relation $\omega(\mathbf{k})$ lies to the right of the light line which means that the surface plasmon have a longer wavevector than light waves of the same energy $\hbar\omega$, propagating along the surface. Therefore they are called "nonradiative" surface plasmons, which describe fluctuations of the surface electron density. Their electromagnetic fields decay exponentially into the space perpendicular to the surface and have their maximum in the surface, as is characteristic for surface waves. The excitation with light needs special light-plasmon couplers (grating couplers, prism couplers), since the dispersion relation lies to the right of the light line.

5.1.3 Dispersion relation for surface plasmons

The electron charge oscillations on a metal boundary can perform coherent fluctuations which are called surface plasma oscillations (144; 145). Their existence has been demonstrated in electron energy-loss experiments by Powell and Swan (146). The frequency ω of these longitudinal oscillations is tied to its wavevector k_x by a dispersion relation $\omega(k_x)$. These charge fluctuations, which can be localized in z direction within the Thomas-Fermi screening length of about 1 Å, are accompanied by a mixed transversal and longitudinal electromagnetic field which disappears at $|z| \to \infty$, Fig. 5.1, and has its maximum in the surface $z = 0$, typical

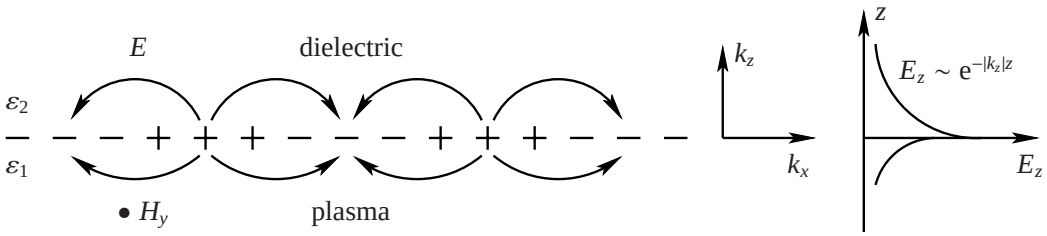

Figure 5.1: The charges and the electromagnetic field of SP's propagating on a surface in the x direction are shown schematically. The exponential dependence of the field E_x is seen on the right. H_y shows the magnetic field in the y direction of this $p-$polarized wave.

for surface waves. This explains their sensitivity to surface properties. The field is described by

$$E = E_0^{\pm} \exp[+i(k_x x \pm k_z z - \omega t)] \tag{5.1}$$

with + for $z \geq 0$, − for $z \leq 0$, and with imaginary k_z, which causes the exponential decay of the field E_z. The wavevector k_x lies parallel to the x direction; $k_z = 2\pi/\lambda_p$, where λ_p is the wavelength of the plasma oscillation. Maxwell's equations yield the retarded dispersion relation for the plane surface of a semiinfinite metal with the dielectric function ($\varepsilon_1 = \varepsilon_1' + i\varepsilon_1''$), adjacent to a medium ε_2 as air or vacuum:

$$D_0 = \frac{k_{z1}}{\varepsilon_1} + \frac{k_{z2}}{\varepsilon_2} = 0 \tag{5.2}$$

together with

$$\varepsilon_i\left(\frac{\omega}{c}\right)^2 = k_x^2 + k_{zi}^2 \tag{5.3}$$

or

$$k_{zi} = \sqrt{\varepsilon_i\left(\frac{\omega}{c}\right)^2 - k_x^2}, \quad i = 1, 2.$$

The wavevector k_x is continuous through the interface. The dispersion relation (5.2) can be written as

$$k_x = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \tag{5.4}$$

If we assume besides a real ω and ε_2 that $\varepsilon_1'' < |\varepsilon_1'|$, we obtain a complex $k_x = k_x' + ik_x''$ with

$$k_x' = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1'\varepsilon_2}{\varepsilon_1' + \varepsilon_2}} \tag{5.5}$$

$$k_x'' = \frac{\omega}{c}\left(\frac{\varepsilon_1'\varepsilon_2}{\varepsilon_1' + \varepsilon_2}\right)^{3/2}\frac{\varepsilon_1''}{2(\varepsilon_1')^2}. \tag{5.6}$$

For real k_x' one needs $\varepsilon_1' < 0$ and $|\varepsilon_1'| > \varepsilon_2$, which can be fulfilled in a metal and also in a doped semiconductor near the eigenfrequency; k_x'' determines the internal absorption. In the following we write k_x in general instead of k_x'.

The dispersion relation, see Fig. 5.2, approaches the light line $\sqrt{\varepsilon_2}\omega/c$ at small k_x, but remains larger than $\sqrt{\varepsilon_2}\omega/c$, so that the SPs cannot transform into light: it is a "nonradiative" SP. At large k_x or

$$\varepsilon_1' \to -\varepsilon_2 \tag{5.7}$$

the value of ω approaches

$$\omega_{sp} = \sqrt{\frac{\omega_p}{1 + \varepsilon_2}} \tag{5.8}$$

for a free electron gas where ω_p is the plasma frequency $\sqrt{4\pi ne^2/m}$, with n the bulk electron density. With increasing ε_2, the value of ω_{sp} is reduced.

At large k_x the group velocity goes to zero, so that the SP resembles a localized fluctuation of the electron plasma.

Spatial extension of the SP fields

Wavevectors k_{z2} and k_{z1} are imaginary due to the relations $\omega/c < k_x$ and $\varepsilon_1' < 0$, see (5.3), so that, as mentioned above, the field amplitude of the SPs decreases exponentially as $\exp(-|k_{zi}||z|)$, normal to the surface. The value of the skin depth at which the field falls to $1/e$, becomes

$$\hat{z} = \frac{1}{|k_{zi}|} \tag{5.9}$$

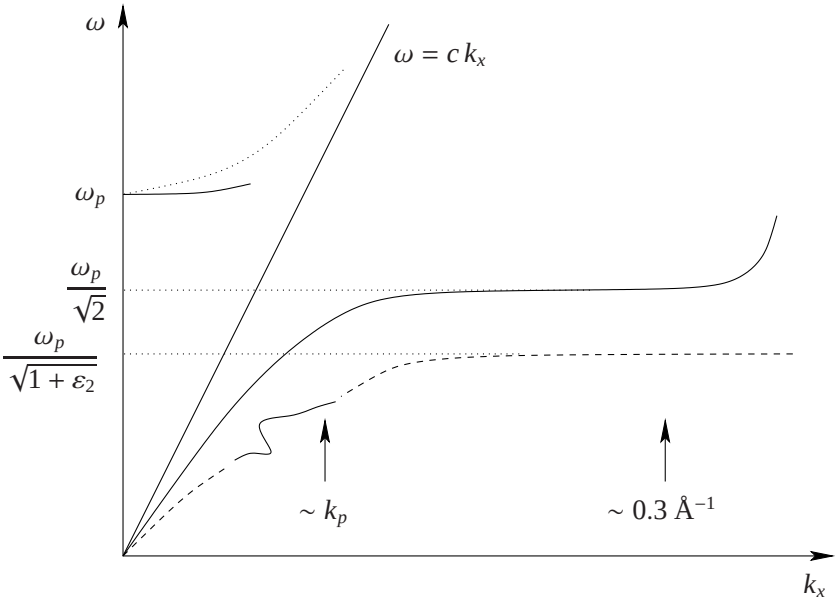

Figure 5.2: The dispersion relation of nonradiative SPs (solid line), right of the light line $\omega = ck_x$; the retardation region extends from $k_x = 0$ up to about $k_p = 2\pi/\lambda_p$ (λ_p plasma wavelength). The dashed line, right of $\omega = ck_x$, represents SPs on a metal surface coated with a dielectric film (ε_2). Left of the light line, $\omega(k_x)$ of the radiative SPs starts at ω_p (solid line). (The dotted line represents the dispersion of light in a metal: $\omega/k_x = c/|\varepsilon_1'|^{1/2}$ of in the case of free electrons $\omega^2 = \omega_p^2 + c^2k_x^2$.) The slight modulation in the dashed dispersion curve comes from an eigenfrequency in a monomolecular dye film deposited on a Langmuir-Blodgett film (ε_2). The lateral depresses $\omega_p/\sqrt{2}$ to $\omega_p/\sqrt{1+\varepsilon_2}$.

or

in the medium with ε_2:

$$\hat{z}_2 = \frac{\lambda}{2\pi}\sqrt{\frac{\varepsilon_1' + \varepsilon_2}{\varepsilon_2^2}}$$

in the medium with ε_1:

$$\hat{z}_1 = \frac{\lambda}{2\pi}\sqrt{\frac{\varepsilon_1' + \varepsilon_2}{\varepsilon_1'^2}}. \tag{5.10}$$

For $\lambda = 600$ nm one obtains for silver $\hat{z}_2 = 390$ nm and $\hat{z}_1 = 24$ nm, and for gold 280 nm and 31 nm, respectively.

At large k_x, $\hat{z}_i$ is given by about $1/k_x$ leading to a strong concentration of the field near the surface in both media.

At low k_x or large $|\varepsilon_1'|$ values, the field in air has a strong (transverse) component E_z compared to the (longitudinal) component E_x, namely $E_z/E_x = -i|\varepsilon_1'|^{1/2}$

and extends far into the air space; it resembles thus a guided photon field (Zenneck-Sommerfeld wave). In the metal E_z is small against E_x since $E_z/E_x = i|\varepsilon_1'|^{-1/2}$. At large k_x both components E_x and E_z become equal: $E_z = \pm iE_x$ (air: $+i$, metal: $-i$).

Propagation length of the SPs

The intensity of SPs propagating along a smooth surface decreases as $e^{-2k_x'' x}$ with k_x'' from (5.6). The length L_p after which the intensity decreases to $1/e$ is then given by

$$L_p = \frac{1}{2k_x''}. \tag{5.11}$$

In the visible region, L_p reaches the value of $L_p = 22\ \mu$m in silver at $\lambda = 5145$ Å and $L_p = 500\ \mu$m (0.05 cm) at $\lambda = 10600$ Å. The absorbed energy heats the film, and can be measured, e.g., with a photoacoustic cell.

Instead of regarding the spatial decay of the SPs along the coordinate x, the temporal decay time T_p can be of interest. The values of L_p and T_p are correlated by $L_p = T_p v_g$ with v_g the group velocity. Assuming a complex $\omega = \omega' - i\omega''$ and real k_x', with $T_p = 2\pi/\omega''$, we obtain from (5.4)

$$\omega'' = k_x' c \frac{\varepsilon_1''}{2(\varepsilon_2')^2} \frac{\varepsilon_1'\varepsilon_2}{\varepsilon_1' + \varepsilon_2}$$

$$\omega' = k_x' c \frac{\varepsilon_1' + \varepsilon_2^{1/2}}{\varepsilon_1'\varepsilon_2}. \tag{5.12}$$

5.1.4 Plasmonic bandgap

In this subsection we consider the influence of periodicity on a plasmon dispersion curve. The propagation of waves in periodic media is fundamental for understanding of many physical phenomena. The properties of photonic materials are based on the interaction between the optical field and the material showing periodicity on the scale of the light wavelength. The periodicity modifies the propagation of the optical wave within the material, and under appropriate circumstances may prohibit propagation over some range of optical frequencies—a photonic band gap.

The photonic materials generally considered are bulk-like in nature. In such systems the photon is dressed by the periodic material—this dressed state is called a polariton mode of the system. One can also consider a system that involves surface rather bulk modes. If the surface is metallic then the relevant mode is SP mode and the periodic perforation can be used to provide the periodicity. Just as in the bulk case, under appropriate conditions this periodicity may result in an energy band gap in the propagation surface modes (147).

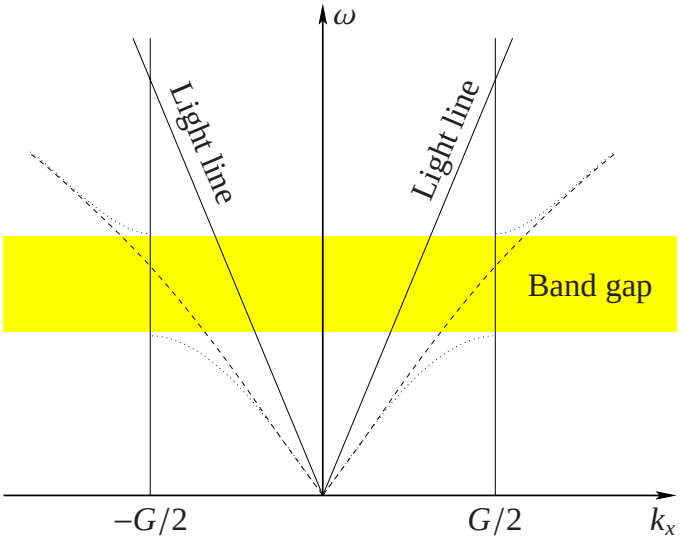

Figure 5.3: On the flat surface there no energy transfer between light and SP is possible.

Such situation is depicted in Fig. 5.3. The SP dispersion curve for a flat (dashed line) and for periodically perforated (dotted line) metal/dielectric interface. The dashed line shows the SP dispersion curve for the flat surface, the dotted line that for a periodically perforated surface. The frequency gap is opened up in the case of the plasmonic crystal. The gap occurs at $\pm G/2$, the zone boundary. The Bragg vector of the plasmonic crystal being $G = 2\pi/a_0$, where a_0 is the lattice constant. Also shown are the light lines. These are the dispersion curves for photons traveling at grazing incidence to the interface between the metal and the dielectric, i.e., those having the largest possible value of k_x: one can see that it is impossible to couple the SP directly to photons; the SP always have more momentum than the photon of the same frequency.

Ban gap formation is not the only effect that is created by periodicity. Because of the grating momentum it is possible to couple light and SP, and hence to experimentally map the SP dispersion curve. Folded SP dispersion curves or plasmonic Brillouin zones shown in Fig. 5.4. In contrast to the flat metal situation, crossing points exist on a perforated surface where the energy exchange is possible. These points are marked with red circles in Fig. 5.4.

In order to map the SP Brillouin zone a zero-order transmission experiment can be used. A sketch of such an experiment are shown in Fig. 5.5. The plasmonic crystal is placed in (x, y) plane. Rotating it around the y-axis allows to study the dispersion relation in the k_x direction. The latter is given by $k_x = |k_0| \sin\theta$, where $|k_0| = 2\pi/\lambda$ is the wavevector of incident radiation; λ is the wavelength. In the transmission spectrum one see different SP branches visualized through grating momentum (see Fig. 5.6).

The interaction between light and SP is made by transferring momentum through the grating:

$$k_x^{sp} = k_x^{light} \pm iG_x \pm iG_y \tag{5.13}$$

where k_x^{sp} is the SP wavevector, k_x^{light} is the component of the incident wavevector

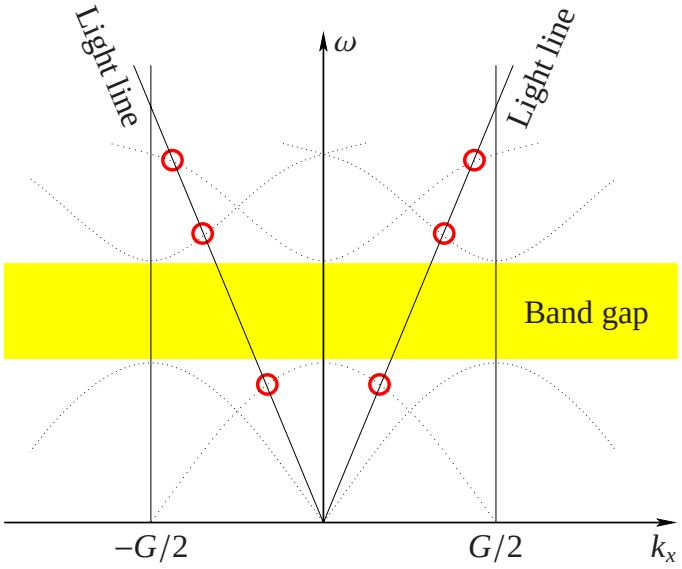

Figure 5.4: The SP Brillouin zone of plasmonic crystal. Energy transfer between light and SP is possible (the coupling points are marked with red circles).

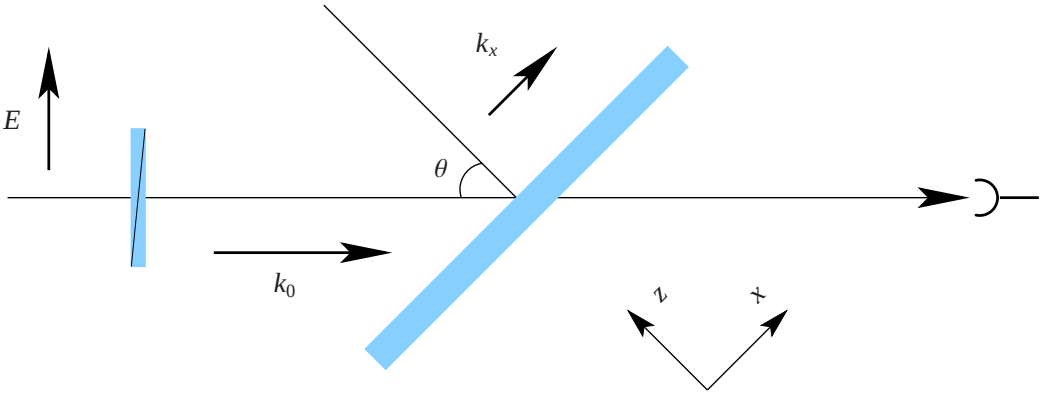

Figure 5.5: Schematic diagram of zero-order transmission experiment.

that lies in the plane of the sample as defined in the Fig. 5.5, G_x and G_y are the reciprocal lattice vectors for a square lattice with $|G_x| = |G_y| = 2\pi/a_0$ and i, j are integers. In order to make a comparison with the data, one use a SP dispersion for a smooth film (5.4), ignoring in a first approximation the fact that the holes in the nodes of plasmonic crystal may cause both a significant change in the plasmon dispersion and a large coupling between the front and back surfaces of the metal film.

At normal incidence $\theta = 0$, Eqs. (5.13) and (5.4) reduce to

$$\lambda \sqrt{i^2 + j^2} = a_0 \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \tag{5.14}$$

From Fig. 5.6 it is clear why the $(\pm 1, 0)$ modes have a fairly large dispersion, while the $(0, \pm 1)$ modes show only weak dispersion since only the x componenet of the wavevector has been changed.

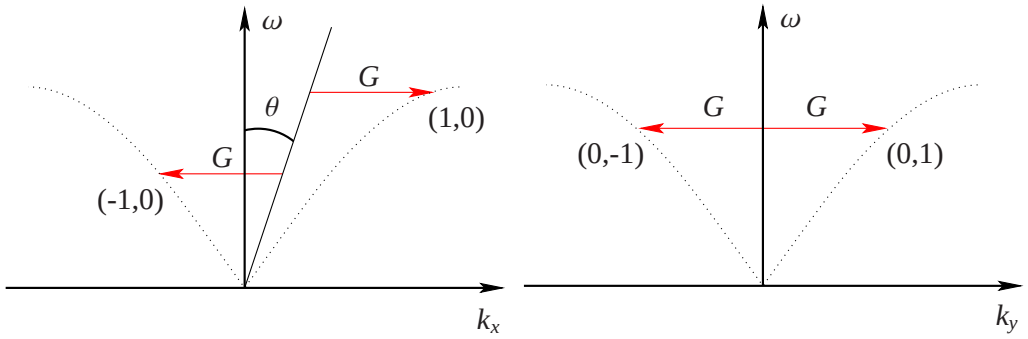

Figure 5.6: Plasmonic crystal's Brillouin zone mapping experiment. Through the grating momentum G one can directly visualize SP dispersion curve. By changing k_x one can see (-1,0), (1,0) and double degenerated (0,-1)/(0,1) SP branches.

A detailed analysis of Brillouin zone mapping experiments is given in subsections 5.4.5 and 5.4.6.

5.1.5 Photonic crystals vs. plasmonic crystals

Fabrication of structures with a tree-dimensional band gap is a challenge because they tend to have complex three-dimensional connectivity and strict alignment requirements (3; 148; 149). Such designs have been the subject of many recent developments (150; 151; 152; 153). An alternative system, the photonic crystal slab, has been proposed that promises easier fabrication using existing technologies (154; 155; 156; 157). This is a dielectric structures that has only two-dimensional periodicity and uses index guiding to confine light in the third dimension. Photonic crystal slabs retain or approximate many of the desirable properties of the true photonic crystal, but at the same time are much more easily realized at submicron length scales (158).

The photonic crystal slab is similar to the two dimensional photonic crystal apart from the fact that the first has a finite thickness. Due to this difference no true band gap exists in slabs. Light confinement in the vertical direction is not through a band gap but through total internal reflection. The band exists only in the direction parallel to slab plane where no guided modes exist. It is not a complete gap because at the same frequencies one can have radiative modes that propagate out of slab plane. The presence of the radiation modes in the gap has the consequence that resonant cavity modes will eventually decay into the background. Such a system being quasi-two-dimensional nevertheless requires more complicated fully tree-dimensional analysis.

Systems similar to photonic crystal slabs are two-dimensional plasmonic crystals. One possible realization is optically thick metal film perforated with a two-dimensional array of holes or other type of scatterers. The interface between the metal and the dielectric medium is the plane of two-dimensional plasmonic crystal. Such systems are significantly easier to produce with the help of, e.g., dry

etching processes. Another advantages is that there is no need in waveguiding for electromagnetic field confinement in the direction perpendicular to the crystal plane. This is possible due to the nature of the surface plasmon mode that has a maximum field intensity on the metal surface and exponentially decay away from it (see section 5.1.3 and appendix B for more details). A limitation of such designs is of course the lossy nature of the metal materials. Damping effects can be limited by shifting the operations frequency into the infrared region, where the plasmon propagation length is much longer.

5.2 Methods of calculations for photonic and plasmonic crystals

5.2.1 Time domain vs. frequency domain

There are two common computational approaches mainly used for studying the linear optical properties of different electromagnetic structures: frequency-domain and time-domain approaches. Each of them has its own unique advantages and disadvantages.

Time-domain methods are well-suited to computing properties that involves the evolution of the fields, such as transmission and resonance decay-time calculations. They can also be used to calculate band gap structures and for finding resonant modes, by looking for peaks in the FT of the system response to some input. The main advantage of this approach is that one obtains the system response in a broad frequency range from a calculation involving the propagation of a single field.

There are several disadvantages to this technique, however. In the FT, the frequency resolution is inversely related to the simulation time; to get 10 times the resolution you must run your simulation 10 times as long. The time-step size must be proportional to the spatial-grid size for numerical stability; thus, if doubling the spatial resolution is necessary, the number of time steps also will double and hence the simulation length will increase by factor of 4. Fourth, you only get the frequencies of the states; to get the eigenstates themselves (so that you can see what the modes look like and do calculations with them), you must run the simulation again.

The frequency-domain are in many ways better-suited to calculating band structures and eigenstates. The error in the frequency in an iterative eigensolver typically decays exponentially with the number of iterations, so the number of iterations is logarithmic in the desired tolerance. The number of iterations typically remains almost constant even if one increase the resolution (the work for each iteration increases, of course). It is possible to get the frequencies and the eigenstates at the same time.

The biggest disadvantage of frequency-domain methods is that one should compute all of the lowest eigenstates, up to the desired one, even if those are not

of interest. This was especially problematic in defect calculations, in which a large supercell is used, because in that case the lower bands are "folded" many times in the Brillouin zone. Thus, it is often needed to compute a large number of bands in order to get to the wanted one (incurring large costs both in time and in storage).

In the next few subsections we will take closer look in particular realization of numerical methods used for studying electromagnetic structures.

5.2.2 Plane wave method

The most commonly used method for studying of band gap structures is the plane wave method (159). This is a frequency domain method that involves solving of Maxwell's equations for a media characterized by a spatially varying dielectric function $\varepsilon(\mathbf{r})$:

$$\nabla \cdot \mathbf{D} = 0 \tag{5.15}$$
$$\nabla \cdot \mathbf{H} = 0 \tag{5.16}$$
$$\nabla \times \mathbf{E} = i(\omega/c)\,\mathbf{H} \tag{5.17}$$
$$\nabla \times \mathbf{H} = -i(\omega/c)\,\mathbf{D} \tag{5.18}$$
$$\mathbf{D}(\mathbf{r}) = \varepsilon(\mathbf{r})\,\mathbf{E}(\mathbf{r}) \tag{5.19}$$

These may be decoupled to generate an equation only in the magnetic field,

$$\nabla \times (\varepsilon^{-1}(\mathbf{r})\,\nabla \times \mathbf{H}) = (\omega/c)^2\,\mathbf{H} \tag{5.20}$$

and only in the electric field

$$\nabla \times (\nabla \times \mathbf{E}) = (\omega/c)^2\,\varepsilon(\mathbf{r})\,\mathbf{E} \tag{5.21}$$

Here we should note that the vector nature of the wave equation has an critical importance. Attempts to adopt the scalar wave approximation (160) led to qualitatively incorrect results. In the simplest case $\varepsilon(\mathbf{r})$ is a real and periodic function of $\mathbf{r}$, it is frequency independent, and the magnetic permeability μ is 1. In this case, the solution scales with the period of $\varepsilon(\mathbf{r})$: reducing the size of the structure by a factor of 2 will not change the spectrum of electromagnetic modes other than scaling all frequencies up by a factor of 2. For more detailed description of the Maxwell equations scaling properties see subsection 2.1.3.

Because of the periodicity of the problem, we can make use of Bloch's theorem to expand the electric and magnetic fields in terms of Bloch waves:

$$\mathbf{H}(\mathbf{r}) = \sum_{\mathbf{K}} \mathbf{H}_{\mathbf{k}} \exp(i\mathbf{K} \cdot \mathbf{r}) \tag{5.22}$$

where $\mathbf{K} = \mathbf{k} + \mathbf{G} \cdot \mathbf{k}$ is a vector in the Brillouin zone and $\mathbf{G}$ is a reciprocal lattice vector.

The solution of the magnetic field has the form of an eigenvalue problem (see subsection 2.2.2)

$$\sum_{\mathbf{K}} \mathbf{K} \times \varepsilon^{-1}_{\mathbf{K},\mathbf{K}'}(\mathbf{K} \times \mathbf{H}_{\mathbf{K}'}) = -(\omega/c)^2 \mathbf{H}_{\mathbf{K}} \tag{5.23}$$

The corresponding equation for the electric field does not have the form of a simple eigenvalue problem since the dielectric function enters into the frequency dependent right-hand side

$$\mathbf{K} \times \mathbf{K} \times \mathbf{E}_{\mathbf{K}} = -(\omega/c)^2 \sum_{\mathbf{K}'} \varepsilon_{\mathbf{K},\mathbf{K}'} \mathbf{E}_{\mathbf{K}'} \tag{5.24}$$

Hence, we obtain the photonic band structure by solving Eq. (5.22) for the magnetic fields.

Here $\varepsilon_{\mathbf{K},\mathbf{K}'} = \varepsilon(\mathbf{G} - \mathbf{G}')$ is the Fourier transform (FT) of the dielectric function. Dielectric functions with sharp spatial discontinuities require an infinite number of plane waves in the Fourier expansion. This problem can be avoided by smearing out of the interfaces of the dielectric objects in the unit cell. For example, for modeling a cylinder of radius a and dielectric ε, one can employ the smeared dielectric function

$$\varepsilon(\mathbf{r}) = 1 + (\varepsilon - 1)/(1 + \exp((r - a)/w)) \tag{5.25}$$

where the width w of the interface is chosen as a small fraction of the radius a (≈ 0.01–$0.05a$). Practically one incorporate the smearing and define the dielectric function $\varepsilon(\mathbf{r})$ over grid in real-space. The FT of the dielectric function in finite plane wave basis set is computed to obtain $\varepsilon(\mathbf{G} - \mathbf{G}')$. The dielectric matrix in Fourier space is then involved to obtain $\varepsilon^{-1}(\mathbf{G} - \mathbf{G}')$. This procedure yields much better convergence than the alternative method of determining $\varepsilon^{-1}(\mathbf{r})$ in real-space and then performing a FT to obtain $\varepsilon^{-1}(\mathbf{G} - \mathbf{G}')$.

The transverse components of the magnetic field are $h_{\mathbf{K},\lambda}$, that is,

$$\mathbf{H}_{\mathbf{K}} = h_{\mathbf{K},1}\mathbf{e}_1 + h_{\mathbf{K},2}\mathbf{e}_2 \tag{5.26}$$

where the unit vectors $\mathbf{e}_1$ and $\mathbf{e}_2$ form an orthogonal triad $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{K})$.

The solution (5.23) for the magnetic field reduces to the eigenvalue problem

$$\sum_{\mathbf{K}',\lambda'} \boldsymbol{M}(\mathbf{K}, \lambda; \mathbf{K}', \lambda') h_{\mathbf{K}',\lambda'} = (\omega/c)^2 h_{\mathbf{K},\lambda} \tag{5.27}$$

The matrix M is defined by

$$\boldsymbol{M}(\mathbf{K}, \lambda; \mathbf{K}', \lambda') = |\mathbf{K}||\mathbf{K}'| \begin{pmatrix} \mathbf{e}_2 \cdot \mathbf{e}'_2 & -\mathbf{e}_2 \cdot \mathbf{e}'_2 \\ -\mathbf{e}_2 \cdot \mathbf{e}'_1 & \mathbf{e}_1 \cdot \mathbf{e}'_1 \end{pmatrix}. \tag{5.28}$$

In practice, the photonic band structure given by the frequencies $\omega(\mathbf{K}, \lambda)$ are computed over several high symmetry points in the Brillouin zone or on a grid in the

Brillouin zone if the density of states is needed. Plane wave convergence is closely checked.

Properties of a photonic band gap depends on (1) the local connectivity of the dielectric structure, (2) the contrast between the two media, and (3) the filling ratio. A minimum dielectric constant ($\varepsilon > 4$) is usually needed to observe the band gaps. The photonic band structure method based on the plane wave approach is a systematic way to search for the existence of band gaps in dielectric structures (2; 161; 162).

5.2.3 Transfer matrix method

While the method described in subsection 5.2.2 focuses on a particular wavevector, there are complementary methods that focus on a single frequency. In the transfer matrix method (TMM), first introduced by Pendry and MacKinnon (163; 164), Eqs. (5.17 and 5.18) are discretized and the z components of the field can be eliminated, so, we derive the following equations:

$$\begin{aligned} E_x(i,j,k+1) &= E_x(i,j,k) + ic\omega\mu_0\mu(i,j,k)H_y(i,j,k) \\ &+ \frac{ic}{a\omega\varepsilon_0\varepsilon(i,j,k)}[a^{-1}\{H_y(i-1,j,k) - H_y(i,j,k)\} \\ &- b^{-1}\{H_x(i,j-1,k) - H_x(i,j,k)\}] \\ &- \frac{ic}{a\omega\varepsilon_0\varepsilon(i+1,j,k)}[a^{-1}\{H_y(i,j,k) - H_y(i+1,j,k)\} \\ &- b^{-1}\{H_x(i+1,j-1,k) - H_x(i+1,j,k)\}] \end{aligned} \tag{5.29}$$

$$\begin{aligned} E_y(i,j,k+1) &= E_y(i,j,k) + ic\omega\mu_0\mu(i,j,k)H_x(i,j,k) \\ &+ \frac{ic}{b\omega\varepsilon_0\varepsilon(i,j,k)}[a^{-1}\{H_y(i-1,j,k) - H_y(i,j,k)\} \\ &- b^{-1}\{H_x(i,j-1,k) - H_x(i,j,k)\}] \\ &- \frac{ic}{b\omega\varepsilon_0\varepsilon(i,j+1,k)}[a^{-1}\{H_y(i-1,j+1,k) - H_y(i,j+1,k)\} \\ &- b^{-1}\{H_x(i,j,k) - H_x(i,j+1,k)\}] \end{aligned} \tag{5.30}$$

$$
\begin{aligned}
H_x(i,j,k+1) &= H_x(i,j,k) - ic\omega\varepsilon_0\varepsilon(i,j,k+1)E_y(i,j,k+1) \\
&+\frac{ic}{a\omega\mu_0\mu(i-1,j,k+1)}[a^{-1}\{E_y(i,j,k+1)-E_y(i-1,j,k+1)\} \\
&-b^{-1}\{E_x(i-1,j+1,k+1)-E_x(i-1,j,k+1)\}] \\
&-\frac{ic}{a\omega\mu_0\mu(i,j,k+1)}[a^{-1}\{E_y(i+1,j,k+1)-E_y(i,j,k+1)\} \\
&-b^{-1}\{E_x(i,j+1,k+1)-E_x(i,j,k+1)\}] \qquad (5.31)
\end{aligned}
$$

$$
\begin{aligned}
H_y(i,j,k+1) &= H_y(i,j,k) + ic\omega\varepsilon_0\varepsilon(i,j,k+1)E_x(i,j,k+1) \\
&+\frac{ic}{a\omega\mu_0\mu(i,j-1,k+1)} \\
&\times[a^{-1}\{E_y(i+1,j-1,k+1)-E_y(i,j-1,k+1)\} \\
&-b^{-1}\{E_x(i,j,k+1)-E_x(i,j-1,k+1)\}] \\
&-\frac{ic}{a\omega\mu_0\mu(i,j,k+1)}[a^{-1}\{E_y(i+1,j,k+1)-E_y(i,j,k+1)\} \\
&-b^{-1}\{E_x(i,j+1,k+1)-E_x(i,j,k+1)\}] \qquad (5.32)
\end{aligned}
$$

Here, $\varepsilon(i,j,k)$ and $\mu(i,j,k)$ are the dielectric constant and the magnetic permeability at the subcell (i,j,k). a, b, c are dimensions of each supercell along the x, y, z directions. Equations (5.29)–(5.32) are connecting the fields at the $k+1$ plane with the fields at the k plane. TMM can be applied for the band calculations of the infinite periodic structure, but the main advantage of this method is that one can calculate transmission and reflection properties of the electromagnetic waves of various frequencies incident on a finite thickness slab of photonic band gap material.

Such calculations are extremely useful in the interpretation of experimental measurement of transmission and reflection data. The TMM method can also be applied to calculate structures containing absorptive and metallic materials.

5.2.4 Finite difference time domain method

The TMM described in previous subsection 5.2.3 is well suited for steady-state solutions. In this subsection we will take look at the finite difference time domain (FDTD) method that is used for general time-dependent solutions including transient behavior. The core of the method is the numerical solution of the Maxwell curl equations:

$$\nabla \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} \qquad (5.33)$$

$$\nabla \times \mathbf{H} = \varepsilon(\mathbf{r})\frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} \qquad (5.34)$$

The derivatives in the Maxwell equations are approximated with finite differences and the electromagnetic field components are located on a Yee cell (165). In the

Yee cell, the E-field components at time $n\Delta t$ are located on the sides of a cube. The magnetic field, H, components at times $(n + 1/2)\Delta t$ are located at the face-centered points of the Yee cell. This results in both spatial and temporal offsets of the two fields when the Maxwell curl equations are solved on each face of the cube. The system is described by a spatial grid. The time step is chosen such that an electromagnetic wave will propagate less than a grid spacing during the time step.

$$E_x^{n+1}(i,j,k) = E_x^n(i,j,k) + \frac{\Delta t}{\varepsilon(i,j,k)}\left[\frac{H_z^{n+1/2}(i,j+1/2,k) - H_z^{n+1/2}(i,j-1/2,k)}{\Delta y} - \frac{H_y^{n+1/2}(i,j,k+1/2) - H_y^{n+1/2}(i,j,k-1/2)}{\Delta z}\right] \tag{5.35}$$

$$E_y^{n+1}(i,j,k) = E_y^n(i,j,k) + \frac{\Delta t}{\varepsilon(i,j,k)}\left[\frac{H_x^{n+1/2}(i,j,k+1/2) - H_x^{n+1/2}(i,j,k-1/2)}{\Delta z} - \frac{H_z^{n+1/2}(i+1/2,j,k) - H_z^{n+1/2}(i-1/2,j,k)}{\Delta x}\right] \tag{5.36}$$

$$E_z^{n+1}(i,j,k) = E_z^n(i,j,k) + \frac{\Delta t}{\varepsilon(i,j,k)}\left[\frac{H_y^{n+1/2}(i+1/2,j,k) - H_y^{n+1/2}(i-1/2,j,k)}{\Delta x} - \frac{H_x^{n+1/2}(i,j+1/2,k) - H_x^{n+1/2}(i,j-1/2,k)}{\Delta y}\right] \tag{5.37}$$

$$H_x^{n+1/2}(i,j,k) = H_x^{n-1/2}(i,j,k) + \frac{\Delta t}{\mu(i,j,k)}\left[\frac{E_y^n(i,j,k+1/2) - E_y^n(i,j,k-1/2)}{\Delta z} - \frac{E_z^n(i,j+1/2,k) - E_z^n(i,j-1/2,k)}{\Delta y}\right] \tag{5.38}$$

$$H_y^{n+1/2}(i,j,k) = H_y^{n-1/2}(i,j,k) + \frac{\Delta t}{\mu(i,j,k)}\left[\frac{E_z^n(i+1/2,j,k) - E_z^n(i-1/2,j,k)}{\Delta x} - \frac{E_x^n(i,j,k+1/2) - E_x^n(i,j,k-1/2)}{\Delta z}\right] \tag{5.39}$$

$$H_z^{n+1/2}(i,j,k) = H_z^{n-1/2}(i,j,k) + \frac{\Delta t}{\mu(i,j,k)}\left[\frac{E_x^n(i,j+1/2,k) - E_x^n(i,j-1/2,k)}{\Delta y} - \frac{E_y^n(i+1/2,j,k) - E_y^n(i-1/2,j,k)}{\Delta x}\right] \tag{5.40}$$

Here, $E_x^n(i,j,k)$ is the x component of the electric field at the n time step in the (i,j,k) subcell.

Arbitrary systems of finite size can be easily modelled by FDTD (166; 167) in order to find their steady-state or transient response. Any combination of metallic and dielectric materials can be treated by FDTD, as well as materials with nonlinear dielectric properties. Common case when FDTD used with a Gaussian pulse source. The fields are numerically intergated to obtain the fields at long times (>1000 time steps). The FT of the scattered and incident fields generates the frequency dependent response of the system. The steady-state response of the system can be calculated for single frequency source fields. Such steady-state calculations should be repeated at desired frequencies.

At the edges of the FDTD cell outer radiation, boundary conditions are frequently employed. Here the incident wave at the boundary is absorbed. Methos to transform the near fields to radiating far fields are then employed. This is particularly necessary for antenna problems where FF radiation patterns are desired. The FDTD method is a very powerful design tool in simulating the electromagnetic response of systems, covering a broad range of frequencies.

The flexibility of the FDTD makes it the method of choice for plasmon calculations presented here (see subsection 5.3.3).

5.3 Surface plasmon nano-optics at near- and far-fields

5.3.1 Sample description and experimental details

Generally, the transmission efficiency of light through subwavelength holes is believed to be exceedingly low (168; 169). When these holes form periodic arrays, however, the transmission efficiency may be enhanced by several orders of magnitude at certain wavelengths (170). This phenomenon has attracted much attention and recent experimental and theoretical work has attributed it to the excitation of SP (171; 172; 144), in terms of which the spectral positions and strengths of the transmission resonances are determined (170; 173; 174; 175; 176; 177; 178; 179; 180; 181; 182; 183; 184; 185; 186; 147; 187). With regard to the microscopic emission patterns, a recent study suggests that, for offresonant excitation, the field intensity is localized around the holes (176). In addition, many theoretical calculations contend that SPs interfere constructively near holes (182; 185; 186).

Figure 5.7(a) shows a scanning electron microscopy image of a typical sample. The FF transmission spectrum of a sample with a 150 nm hole diameter and a period of 850 nm in a 300 nm thick gold film is displayed in Fig. 5.7(b), with light incident on the sapphire side. It shows transmission peaks at SP resonances corresponding to surface charge oscillations at the air–metal and sapphire–metal interfaces, respectively. The 930 nm peak corresponds to the air metal SP resonance of the first order with its symmetry along the x or y axes defined as horizontal and vertical, respectively. These are termed the [1,0] or [0,1] modes. The 1100 nm peak is the [1,1] sapphire–metal SP resonance of second order with its symmetry

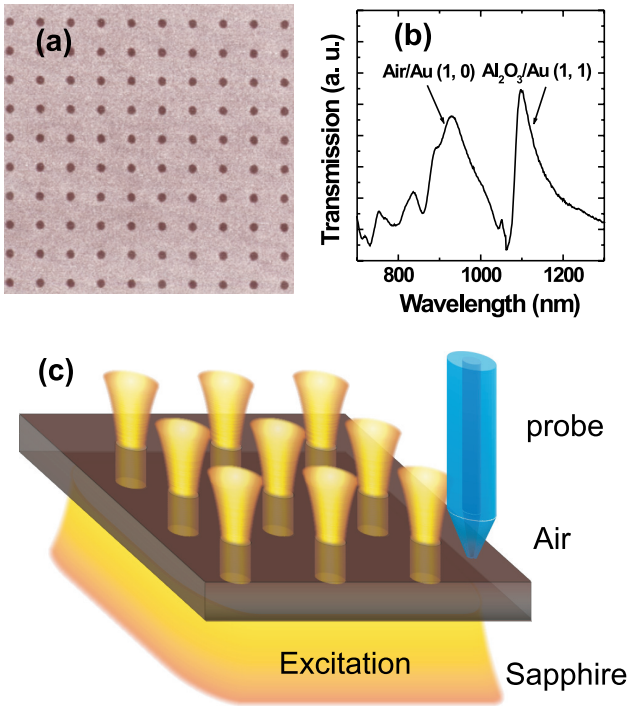

Figure 5.7: (Color) (a) Scanning electron microscope image of a typical periodic array of holes in a thick metal film. (b) FF transmission spectrum. Holes with 150 nm diameters were made on 300 nm thick gold films with a period of 850 nm. (c) Schematic diagram of the experiment.

along the axis diagonal. We employ the NF microscope (see detailed discription in section 3.3) in the transmission geometry where a Ti:sapphire laser excites the sample on the sapphire metal side near the normal incidence and a metal-coated fiber tip with a sub-100 nm aperture collects light on the air–metal side. The incident beam is focused to a spot 5–10 μm in size on the sample. Figure 5.7(c) is a schematic sketch of light emission from the holes, as our macroscopic experience implies.

5.3.2 Light emission from the shadows

Figure 5.8 shows a NF emission pattern for each polarization with the wavelength of an incident beam around the air metal [1,0] SP resonance peak as shown in Fig. 5.7(b). While the transmission efficiency in the FF region is independent of polarization (176), the NF images show a strong polarization dependence. We observe stripes that run essentially perpendicular to the polarization direction. In addition, there are minima between the stripes, so that a peak-to-valley ratio of 10 or larger is seen. The polarization dependences are consistent with the fact that the propagation direction of the SP waves is parallel to the exciting light polarization and that SP waves are mostly longitudinal. In all images most of the light transmitted is found on metal surfaces away from the holes, so that over 90% of the integrated light emission originates from the metal surfaces. This behavior is different from previous observations (176) and recent theoretical predictions (182).

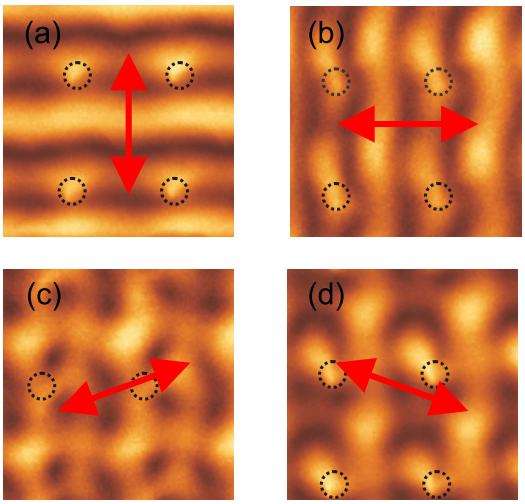

Figure 5.8: (Color) Polarization dependence of NF transmission images near the air-gold [1,0] and/or [0,1] resonance using light with wavelength $\lambda = 877$ nm. The data were taken from the same sample as in Fig. 5.7(b). The arrow represents the polarization of the incident light.

It is important to study which part of the NF pattern can reach the FF where enhanced transmission is observed. We have varied the distance z between the tip and the surface, at a wavelength that is on the high-energy side of the air–metal [1,0] resonance (877 nm). As Figs. 5.9(a)–5.9(d) show, the pattern is preserved well up to $z \simeq 2.2\,\mu$m, which is well over twice the wavelength. In fact, we can observe vertical stripes up to a distance of 5 μm, when the wavelength of the incident light is close to the lattice constant. This clearly indicates that a large portion of the field emanating from nonilluminated metal surfaces is not evanescent: metal shadows can indeed shine into the FF. Figure 5.9(e) plots the total intensity, spatially integrated over the xy plane, as a function of z. It shows clearly that the light transmitted has both evanescent and nonevanescent components. The evanescent part constitutes about half the total integrated intensity with the decay length of the order of 200 nm.

Figure 5.9(f) reveals how simple the emission pattern becomes at large z: the cross-sectional intensity of Fig. 5.9(d) can be fit nearly exactly by $A \pm B \sin(2\pi x/a_0)$ where a_0 is the lattice constant, and A and B fitting parameters. The fact that the region of near-constant integrated intensity of $z \geq 0.5\,\mu$m in Fig. 5.9(e) coincides with the simple patterns of Figs. 5.9(c) and 5.9(d) is extremely instructive. In the NF, many in-plane Fourier components, $k_{||} = 0, \pm 2\pi/a_0, \pm 4\pi/a_0$, contribute. That is why the spatial pattern is generally complicated. As z becomes larger, the larger wavevector modes with small penetration depth into the air side all quickly decay out, and only the $k_{||} = 0, \pm 2\pi/a_0$ modes can survive.

5.3.3 FDTD simulation

In order to better understand our results, we have performed three-dimensional FDTD (see section 5.2.4) simulations of light transmission through 100 nm holes in a free-standing gold film of 100 nm thickness (Fig. 5.10). The field distributions are calculated for linear y polarization of the incident field by numerically

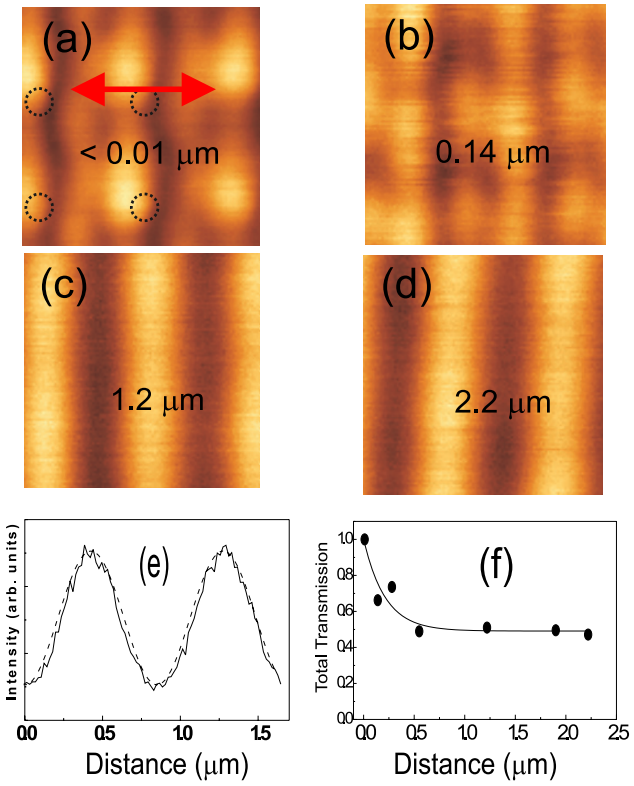

Figure 5.9: (Color) Tip–sample distance (z) dependence of transmission images near the air-gold [1,0] resonance using $\lambda = 877$ nm. The sample is the same as in Fig. 5.8. (a) $z \leq 0.01$ μm. (b) $z \simeq 0.14$ μm. (c) $z \simeq 1.2$ μm. (d) $z \simeq 2.2$ μm. (e) Horizontal cross section of (d) (solid line) and the sinusoidal fit (dotted lines). (f) Integrated total transmission as a function of z.

integrating Maxwell equations in the time-domain assuming a Drude model for the dielectric function of the metal. For simplicity we chose a lattice constant of the hole array of 500 nm and an excitation wavelength of 530 nm, similar to the experimental conditions. These simulations show [Fig. 5.10(a)] that the field distribution close to the aperture is described rather well by the Bethe–Boukamp model, and show typical divergence at the rim of the aperture (168; 182). In addition, strong SP fields are generated on the nonilluminated side and the stripe-like patterns run perpendicular to the polarization direction. Interference of the SP field is clearly visible in the local enhancement of the field intensity along lines in the center between adjacent holes, perpendicular to the incident y polarization, as observed experimentally [Fig. 5.8(a)]. A cross section $|E_y(x = a/2, y, z)|^2$ in the y–z plane along a plane ($x = a/2$) through the center of the holes is given in Fig. 5.10(b). The z dependence of the field emitted from the holes and the SP field on the metal surfaces are different. The emission from the holes decays rapidly with an increase in z and it appears that the field that originates from the metal can dominate the FF emission. The numerical complexity makes it difficult to extend these calculations to the FF region. Yet, two-dimensional simulations suggest similar NF distributions and show indeed a sinusoidally modulated FF emission that stems from the metal surfaces. The main discrepancy between theory and the experiment is that, in theory, holes are still the brightest spots, whereas in experiments, that is not the case.

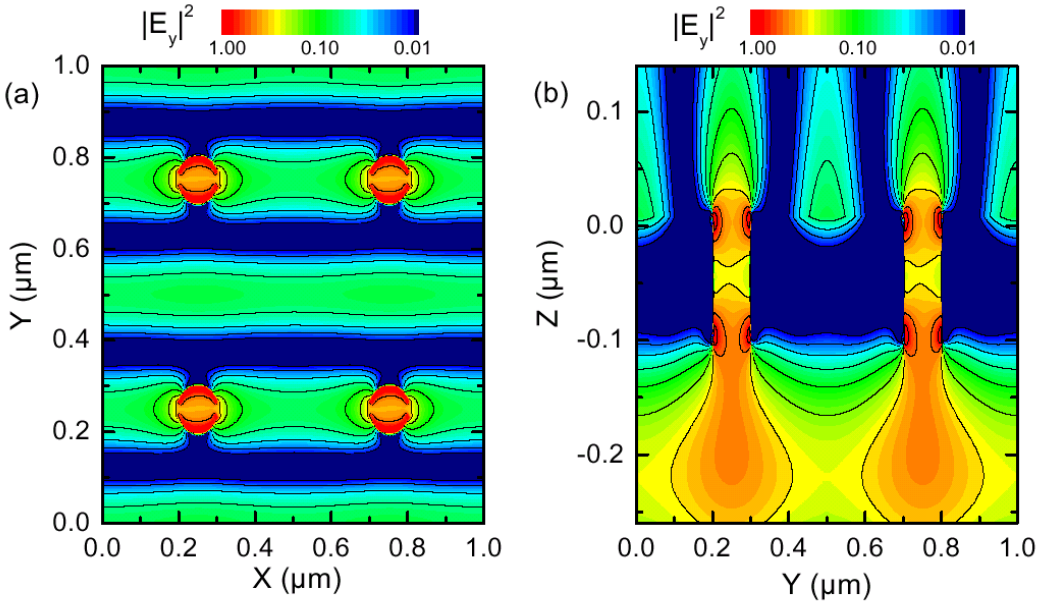

Figure 5.10: (Color) Calculated normalized NF intensity $|E_y|^2$ along one hole array period (on a logarithmic gray scale). (a) $|E_y(x, y, 0)|^2$ in the x–y plane showing the SP field on top of the metal film ($z = 0$); (b) cross section $|E_y(x = a_0/2, y, z)|^2$ in the y–z plane on top of the metal film. The cross section is taken along a line ($x = a_0/2$) through the center of the holes.

The comparison between experiment and theoretical simulation clearly indicates that mostly the transversal E_y component of the electrical field parallel to the incident field is coupled into the NF fiber. Coupling of the E_z component into the metalized fiber is strongly suppressed.

5.4 Microscopic Origin of Surface Plasmon Radiation in Plasmonic Crystalls

5.4.1 Problems, samples and used techniques

The metal films perforated with subwavelength hole arrays show unusually high transmission of light at SP resonances (170; 174; 178; 180; 182; 183). Phenomenologically, the incident light is grating coupled to SP excitations on the metal interfaces. SPs on either side of the metal film are coupled through the nano-holes and may eventually be re-emitted into FF radiation. This enhances the on-resonance transmission by 2-3 orders of magnitude over that of isolated nano-apertures, giving rise to a wide range of possible applications, e.g. as NF sources, in photonic integrated circuits or in lithography (170; 188).

These optical properties can be described by eigenvalue solutions (see section

2.2) of Maxwell equations in a periodic dielectric medium (189; 2). The real and imaginary part of the eigenvalues reflect the energy positions of the transmission resonances and the dephasing of the Bloch eigenstates, i.e. the eigenvectors, respectively. The spectral resonances have been investigated (174) and the energy gaps in the dispersion relation were observed (170). Much less is known about the electric field profiles and about the damping properties of SP excitations in nano-hole arrays, i.e. about the microscopic origin of the line shapes of the transmission spectra. Such information is of vital importance for a microscopic understanding of the underlying physics as well as for optimization of such nanostructures for possible applications. Here, NF optical techniques are powerful tools, as they allow for a direct mapping of the eigenmodes of subwavelength structures (190) and their damping rates.

Nano-hole arrays are fabricated by a dry etching technique after e-beam patterning on a 300 nm thick gold film grown on a $\lambda/5$ flat sapphire substrate.

The optical axis of the sapphire substrate is perpendicular to the metal film plane. The ordinary and extraordinary refractive indices of sapphire at 800 nm are $n_o = 1.760$ and $n_e = 1.752$, respectively. The incident light is focused onto the sample with a numerical aperture of $NA = 0.1$, i.e. under a maximum angle of $\theta = 5°$. From the index ellipsoid we can then estimate the effective refractive index for the extraordinary ray propagating under $\theta/2 = 2.5°$. as $n_{eff} = \sqrt{cos(\theta)^2/n_0^2 - sin(\theta)^2/n_e^2}) = n_o + 1.5 * 10^{-5}$. The phase difference between ordinary and extraordinary ray at the exit of the $d = 0.2$ mm thick sapphire substrate is hence negligible. This is experimentally verified by probing the outgoing polarization, after having passed through the sapphire plate. The polarization rotation is indeed negligible.

The damping experiments are performed on a sample with a hole radius r of about 125 nm and an array period a_0 of 761 nm. Its FF emission spectrum [Fig. 5.11(a)] shows transmission peaks assigned to SP resonances at either the

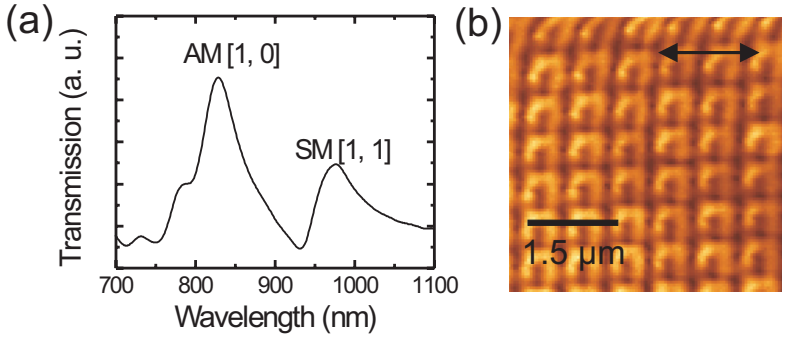

Figure 5.11: (a) Transmission spectrum a gold nano-hole array with $a_0 = 761$ nm and $r = 125$ nm. (b) NF transmission image at an excitation wavelength of 820 nm.

air-metal (AM) or sapphire-metal (SM) interfaces (174). The peak at 825 nm cor-

responds to the AM [1,0] mode, and that at 980 nm is the SM [1,1] SP resonance with its symmetry along the grating diagonal. We use a NSOM in transmission geometry. The sample is illuminated with linearly polarized light from a Ti:sapphire laser through the substrate and the electric field intensity at the metal-air interface is coupled into a metal-coated NF probe with sub-100 nm aperture diameter.

5.4.2 Surface plasma damping experiments

Fig. 5.11 (b) shows the NF emission pattern at an excitation wavelength of 820 near the AM [1,0] resonance, within a scan range of about $5 \times 5\ \mu\text{m}^2$. The emission pattern is dominated by stripe-like patterns that run perpendicular to the polarization direction, indicated by an arrow. The pattern repeats itself with the lattice periodicity a_0, and reflects the field intensity of the excited Bloch eigenstates of the plasmonic band gap system at the metal-air interface. It is given as a coherent superposition of SP plane waves with wavevectors $\pm\frac{2\pi}{a_0}(m,n)$, where m and n are integers. FT show that only a few of the lowest order Bloch waves $(m,n) \leq 2$ are strongly exited for samples with hole radii of more than 100 nm. Generally, all vector components of the local electric field may be detected by an NSOM probe and the image is a superposition of $|E_x|^2, |E_y|^2$, and $|E_z|^2$ (190). The used metal coated glass fibers are insensitive to $|E_z|^2$ due to the boundary conditions at the air-glass interface. Mostly the field component along the incident polarization is detected. This is verified by the quantitative agreement between experimental NF images and three-dimensional FDTD simulations of $|E_x|^2$ (191) (see section 5.3.3), whereas solutions for $|E_z|^2$ are markedly different. These field profiles show a pronounced wavelength dependence and depend strongly on whether AM or SM resonances are excited, as recently predicted (183) for such quasi-two-dimensional band gap structures. Here, it is important that these experiments verify directly that SP modes are efficiently excited at the metal-air interface and that these modes are probed with high spatial resolution with our NF microscope setup.

To measure the damping of SP excitations, we first probe their spatial propagation length, employing a NF microscope (see section 3.3) in the experimental geometry shown in Fig. 5.12(a).

The AM [1,0] SP mode is resonantly excited at the metal-air interface of a first nanohole array. This mode is scattered at the edge of the array, where only the $k_x = 2\pi/a_0$ component of this eigenmode is allowed to propagate onto the flat gold surface. Here its propagation length is on the order 40 μm (192). To measure its damping inside a nano-hole array, we probe the decay of the detected field intensity I_{SP} that occurs when this mode encounters a second hole pattern. This experiment is sensitive to both SP population decay (characteristic time T_1) and purely phase disturbing scattering processes (characteristic time T_2^* (193)). Population decay leads to a spatial decay of the field intensity, whereas pure dephasing processes would lead to a change in the shape of the NF pattern, specifically its image contrast. In a relaxation-time approximation, the total damping value is

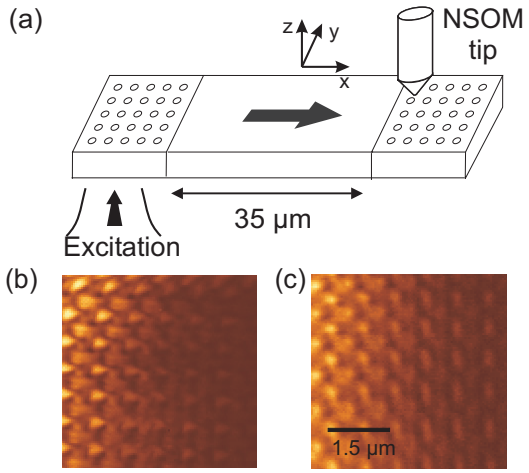

Figure 5.12: (a) Schematics of the experiment. The incident polarization is along x, the SP propagation direction (arrow). (b), (c) NF images of the intensity $I_{SP}(x,y)$ on the non-illuminated grating side with a scan area of $5 \times 5\ \mu m^2$. The excitation was at 760 nm (b) and 830 nm (c), close to the AM [1,0] resonance (790 nm).

given as $1/T_2 = 1/T_1 + 1/T_2^*$, where T_1 is a population decay time and T_2^* is the time constant for pure dephasing. Therefore we refer to the dephasing time T_2 as the time determines the homogeneous linewidth of the system, no matter whether inelastic damping or pure dephasing is involved. Shown in Fig. 5.12(b) and 5.12(c) are $5 \times 5\ \mu m^2$ scans of I_{SP} at excitation wavelengths of 760 and 830 nm, respectively. Both images reveal periodic pattern formation through interference of Bloch waves generated through SP scattering at the holes. As we probe deeper into the grating, the spatial patterns remains unchanged but its overall intensity shows a rapid decay decreasing by more than an order of magnitude within 5 micron. This shows that there is a strong *inelastic* SP damping on a length scale that is much shorter than that in flat metal surfaces. Such an experiment is sensitive to the both inelastic and quasi-elastic scattering events. Inelastic scattering leads to a spatial decay of the field intensity, as observed experimentally. Pure dephasing due to phase fluctuations would lead to a change in the shape of the NF pattern, specifically its image contrast. Such a change is not observed. Therefore, the measurements shows that SP damping due to inelastic scattering into FF radiation is dominant and pure dephasing processes are of minor importance ($T_2^* \gg T_1$). Shown in Fig. 5.13(a) are vertically-integrated intensity profiles $\int I_{SP}$ of the emission patterns at excitation wavelengths 760 and 830 nm. Both curves show, on average, an exponential decay with decay constants of 1.34 and 1.63 μm, respectively. We define the propagation length L_p as the distance over which SP field decays by 1/e, i.e. I_{SP} decays as $I_{SP} \propto \exp(-2 \cdot x/L_p)$. We deduce propagation lengths of $L_p = 2.68$ and $3.26\ \mu$m.

For propagation of a plane wave through a dissipative medium, the damping time T_{2p} of the wave and the coherent propagation length L_p are linked through $T_{2p} = L_p/v_{ph}$, where v_{ph} denotes the SP phase velocity, which is nearly the speed of light (144; 194). In our sample, a_0 is 761 nm and the AM [1,0] mode has its peak at 827 nm. Therefore, v_{ph} can be estimated as $\frac{761}{827} \cdot c \simeq 0.92 \cdot c$, where c is the speed of light. We then deduce SP damping times T_{2p} of 9.7 fs and 11.8 fs, respectively.

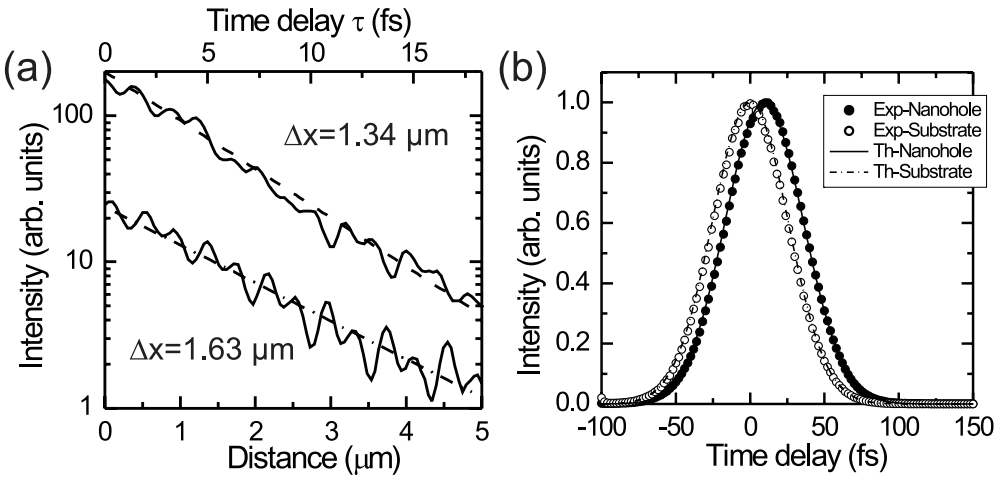

Figure 5.13: (a) Decay of $\int I_{SP}(x,y)dy$ of Fig. 5.12 (b,c), versus distance x. The excitation was at 760 nm (top) and 830 nm (bottom). Dashed line: exponential decay. (b) Cross-correlations of fs pulse transmission (λ = 790 nm) through a nano-hole array (filled circles) and through the substrate only (open circles, performed on a non-metalized piece of the same sample). The uncertainty in time-delay due to birefringence and thickness variations is < 1 fs (uncertainty in optical path length is < 200 nm). Solid line: Curve calculated from Eq. (5.41) assuming T_2 = 10 fs.

5.4.3 Time domain experiments

To confirm the equivalence between spatial propagation and temporal damping of the SP excitation, we perform a direct measurement of the SP damping in the time domain. Here, we have a simple physical picture in mind. The incident pulse $E_{in}(t)$ drives the AM [1,0] mode polarization P_{SP} at its eigenfrequency ω_{SP}. The finite damping of the polarization with T_2 gives rise to a delay in the reemission of the SP field, described by a driven harmonic oscillator equation

$$\frac{d^2P_{SP}}{dt^2} + \frac{2}{T_2}\frac{dP_{SP}}{dt} + \omega_{SP}^2 P_{SP} \propto E_{in}(t). \tag{5.41}$$

All transmitted light is assumed to originate from SP radiation. Assuming $T_2^* \gg T_1$, the transmitted intensity $I_{tr}(t)$ is directly proportional to $| P_{SP}(t) |^2$.

We performed the non-collinear second order intensity cross-correlation experiment. The 40 fs laser pulses passing through the nano-hole sample is overlapped in a BBO crystal with a replica of the incident laser pulses. The intensity of the second harmonic from the BBO is detected. As it is a non-collinear cross-correlation experiment, it is insensitive to the phase of laser pulses. The substrate thickness varies by less than $\lambda/5$, and therefore, the error from substrate thickness

variations is less than 1 fs. The laser is centered at 790 nm, close to the AM [1,0] resonance.

Fig. 5.13(b) (top curve) compares cross-correlations recorded for transmission through the nano-hole array (solid line) and through the substrate only (dotted line). A clear time delay of 10 ± 1 fs is observed, when the pulse is transmitted through the nano-hole structure. In the bottom part of Fig. 5.13(b), the experimental results are compared to a simulation based on Eq. (5.41). Here, a Gaussian profile is taken for the 40 fs input pulse. Good agreement with experiment is obtained if a SP damping time T_{2d} = 10 fs is assumed. This damping time T_{2d} = 10 fs of the total SP polarization matches precisely the value of T_{2p} = 9.7 fs found for an excitation wavelength of 760 nm in the propagation experiments. This indicates strongly that it is indeed the finite damping of the driven SP resonance that is responsible for the delay in light transmission[1].

We now relate the measured damping times to the lineshape of the SP resonances in the optical FF transmission spectra. These spectra are recorded with a spatial resolution of 1 mm, i.e., much larger than the coherent SP propagation length L_p of few μm. Hence, both homogeneous and inhomogeneous broadening may contribute to the FF lineshape, as, e.g., in the optical spectra of systems such as thin semiconductor QWs or ensembles of metal nanoparticles. In the absence of inhomogeneous broadening the homogeneous linewidth of a single Lorentzian resonance is $\Gamma = 2\hbar/T_2$, where T_2 is the total dephasing time. In our sample, the AM [1,0] mode has a FWHM of Γ = 140 meV, corresponding to T_2 = 9.4 fs. From the measured T_{2p} = 9.7 and 11.8 fs, we calculate linewidths of 135 and 112 meV, respectively. Therefore, SP damping can account for most of the measured linewidth in the FF transmission spectra. Contributions from inhomogeneous broadening, arising from structural imperfections of the nano-hole array, are of minor importance and the system can be considered to be *homogeneously broadened*. This equivalence between spectral linewidth and independently measured T_2 has been confirmed for different resonances and samples.

5.4.4 Microscopic origin of the SP damping

Having established that the system is homogeneously broadened, we now discuss the microscopic origin of the SP damping. The propagation experiments suggest that Γ is mostly limited by a *radiative* SP decay into light through scattering at the nano-holes. Such scattering should well be described by classical Mie theory in the Rayleigh limit $a_0 \ll \lambda$. For Rayleigh scattering of 3D waves by small spheres the pronounced wavelength and radius dependence of the scattering cross section $\sigma \propto r^6/\lambda^4$ is well known (196). For scattering of quasi-2D surface bound waves by single circular apertures, an r^4 dependence of σ has been predicted, as expected for dipole scattering (197).

[1] In a separate experiment (195), a time delay of about 6 fs was measured using 100 fs pulses. This delay was tentatively assigned to the reduced group velocity inside the apertures.

To show that the SP damping is indeed dominated by Rayleigh scattering, we first study the wavelength dependence of the spectral linewidths. Fig. 5.14(a) plots, on a double logarithmic scale, the widths Γ of the individual SP resonances

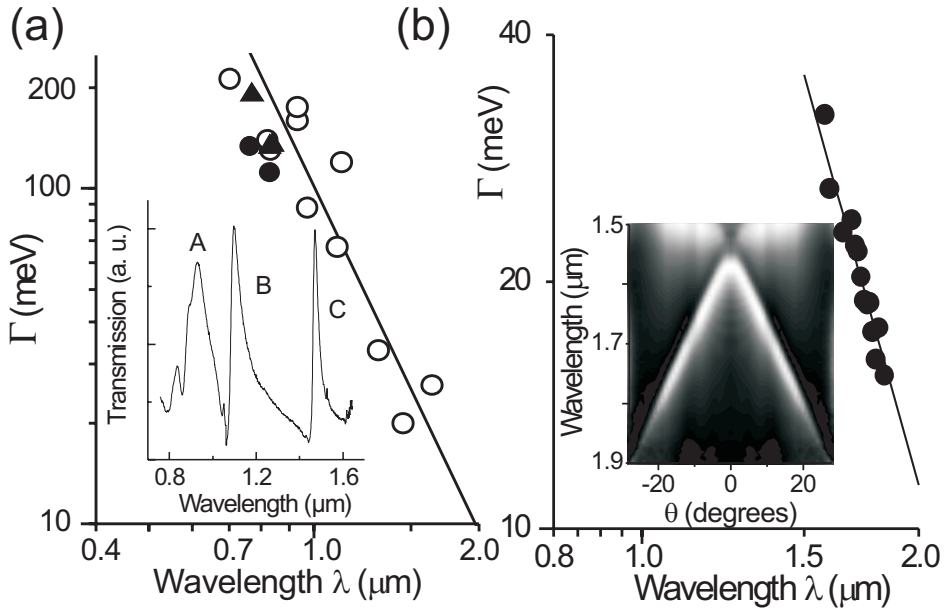

Figure 5.14: (a) Log-log plot of line widths Γ versus peak wavelengths λ for different nano-hole arrays with $r = 125$ nm. The solid line is a fit of Γ to λ^{-n}, with $n = 3.4$. Inset: Transmission spectrum for a gold sample with $a_0 = 800$ nm and $r = 125$ nm. A, B, and C correspond to AM [1,0], SM [1,1], and SM [1,0] resonances. (b) Γ vs. λ for various incident angles for the SM [1,0] peak and fit to λ^{-4} (line). Angle-dependent transmission spectra of the SM [1,0] resonance.

of four samples with similar hole radii (open circles). Widths deduced from time-delay and propagation measurements are shown as filled triangles and filled circles, respectively. We find that Γ scales as λ^n with $n = -3.4 \pm 0.5$. This is indicative of Rayleigh scattering, yet the data show considerable scatter, making a final judgment difficult. This scatter is not surprising, as the damping of *different* AM and SM SP eigenmodes is compared. Mode dependent scattering amplitudes give rise to fluctuations in Γ among the different resonances. To avoid such ambiguities, we probe the wavelength-dependent linewidth of a *single* resonance by measuring angle-dependent transmission spectra near the SM [1,0] resonance. The Γ values in Fig. 5.14 are taken as the FWHM of the SP resonances in the FF spectra. The λ^{-4} dependence of Γ is confirmed by a full lineshape analysis described in next subsection (5.4.5). This analysis gives evidence for an energy gap around $k_x = 0$ (see subsection 5.4.6). In the inset Fig. 5.14(b), the normalized intensity is plotted as a function of λ and incident angle θ. The shift of the resonance position with θ is well described by the model outlined in (174) and an energy gap is clearly observed around $\theta = 0$. Fig. 5.14(b) clearly shows the power-law scaling of Γ and

we now find a slope $n = 4.2 \pm 0.3$, unambiguously demonstrating that the scattering mechanism is indeed Rayleigh-like. Note that the narrow linewidth of only 15 meV around 1.8 μm corresponds to a T_2 of 90 fs and a coherent propagation length of $L_p = 25\,\mu$m, i.e., about 30 lattice periods.

5.4.5 Determination of the FWHM of the transmission peaks

We have extracted the FWHM of the transmission peaks in two different ways. First, we have simply taken the FWHM from the spectra and plotted them as a function of k-vector. These values are given in Fig. 5.14(b). Second, we have

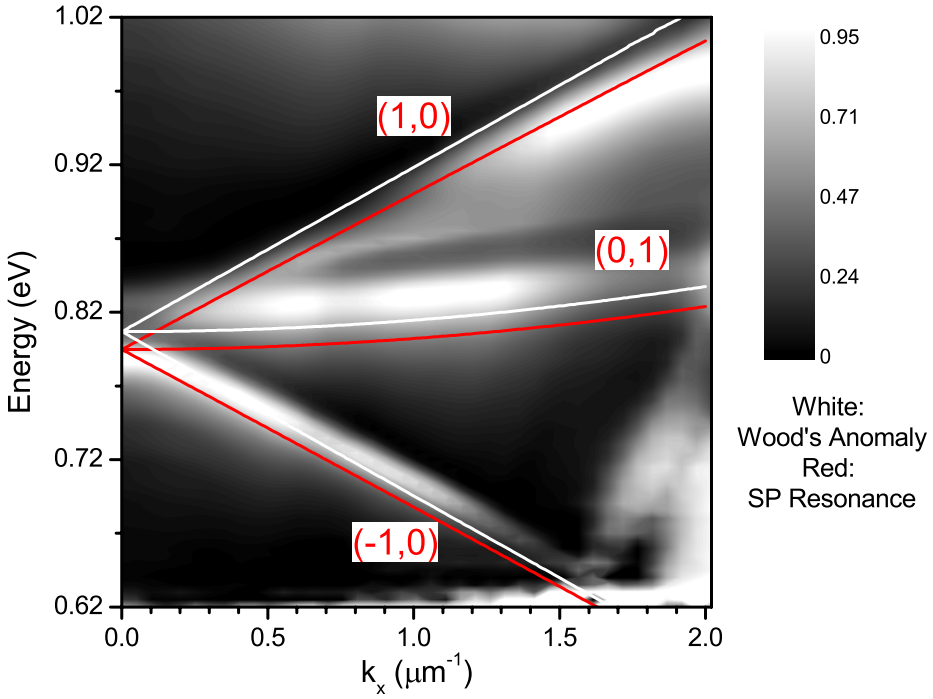

Figure 5.15: (a) Angle-dependent transmission spectra of the periodic nanohole array.

performed a more detailed lineshape analysis of the angle-dependent transmission spectra shown in the inset of Fig. 5.14(b). The full data set in the energy range from 0.62 to 1.02 eV is shown in Fig. 5.15 below. The assignment of the SM[+/-1,0] and SM[0,1] resonances follows Ghaemi et al. (174). The solid red and white lines show simulations of the dispersion relations for the SP resonances and Wood's anomalies similar to those given by Ghaemi et al. Tabulated values for the dielectric function of gold are used.

The transmission spectra at a fixed value of k_x were then fit to a sum of three to four Lorentzian oscillators. (In our spectra, we noted a splitting of the SM[1,0] transmission resonance for $k_x > 0.5\mu\text{m}^{-1}$. In this range of k-vectors, it was necessary to fit the SM[1,0] resonance to a sum of two Lorentzian oscillators in order

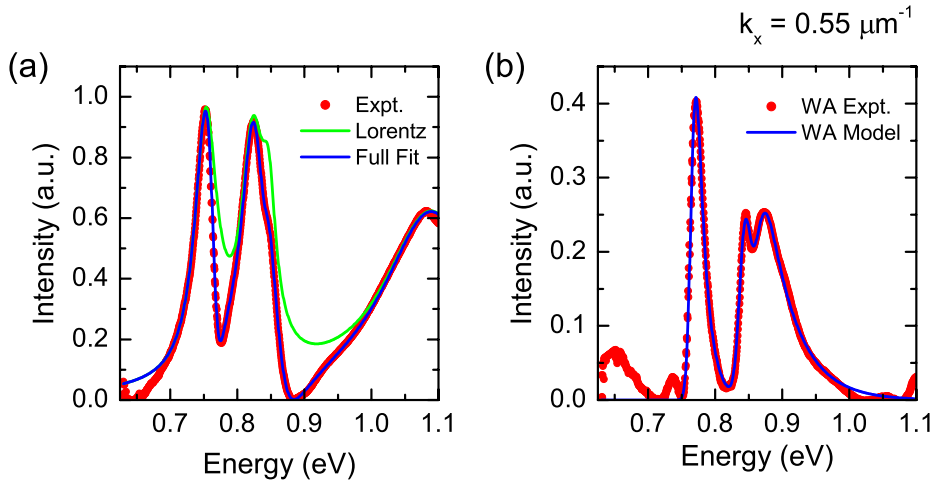

Figure 5.16: (a) Transmission spectrum for $k_x = 0.55\ \mu\text{m}^{-1}$. Circles: Experiment. Solid green line: Fit to a sum of 4 Lorentzian oscillators. Solid blue line: Fit to the full lineshape model. (b) Circles: Experimental lineshape of Wood's anomaly, taken as the difference between Lorentzian oscillator fit and experimental data. Solid line: Fit to Wood's anomaly as described in the text.

to get sufficient agreement with experiment.) These Lorentzian oscillator fits are shown in Fig. 5.16 and 5.17 for two representative transmission spectra at $k_x = 0.55\ \mu\text{m}^{-1}$ and $k_x = 0.76\ \mu\text{m}^{-1}$. It is evident, that the Lorentzian oscillator model gives a good fit of the low energy side of each transmission resonance (i.e. the SP resonance) but fails to account for the asymmetry of transmission spectrum. This asymmetry is generally assigned to the so called Wood's anomaly. Wood's anomalies are well known minima in the reflection spectrum of diffraction gratings. Their spectral lineshape has been the focus of a substantial number of experimental and theoretical investigations (see, e.g. (198; 199; 200)).

We follow the paper by Ghaemi et al. (174) and assume that we can separate the ransmission resonance into the enhanced SP peak and a minimum due to Wood's anomaly. We fit the enhanced SP resonance to a Lorentzian of width γ. We then extract the spectrum of Wood's anomaly by taking the difference between the measured spectrum and the Lorentzian fit. The obtained difference spectra are plotted in Fig. 5.16(b) and 5.17(b).

We find spectrally sharp asymmetric resonances for each of the Wood's anomalies associated to the SM[+/-1,0] and SM[0,1] resonances. For all spectra that we have investigated (about 50), their spectral lineshape is well fit by the convolution of a Gaussian

$$G(E) = \exp(-\beta E^2) \tag{5.42}$$

with a single-sided exponential

$$D(E) = \begin{cases} 0 & \text{if } E < E_W \\ \exp(-\gamma_W (E - E_W)) & \text{if } E \geq E_W. \end{cases} \tag{5.43}$$

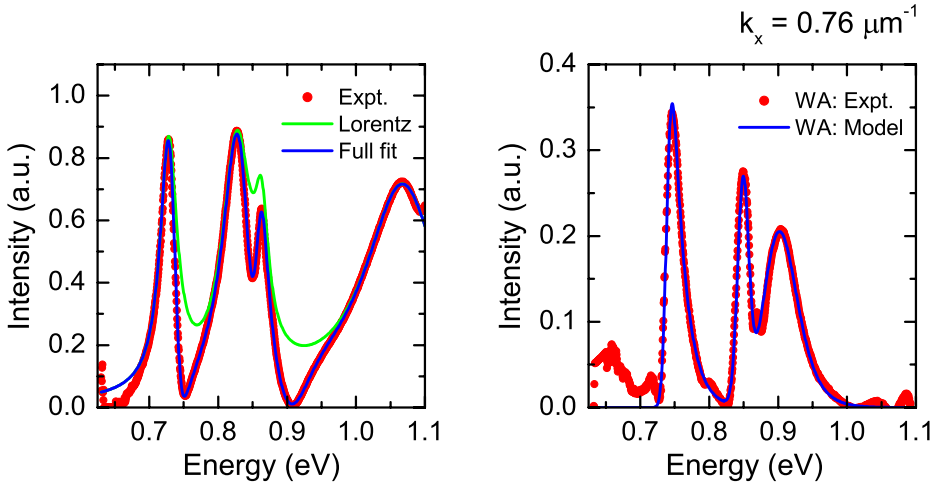

Figure 5.17: The same as in Fig. 5.16 but for $k_x = 0.76\,\mu\text{m}^{-1}$.

The excellent fit of both the difference spectrum and the total transmission spectrum to this model is clearly seen in Fig. 5.16 and 5.17. Such a fit was performed for each value of k_x and the extracted values of the linewidth γ of the Lorentzian fit to the SM[-1,0] resonance is plotted in Fig. 5.18. The data are shown together with the FWHM of the transmission resonance (Fig. 5.16(b), open circles). They are plotted on a double logarithmic scale and the solid lines show fits to the data to a λ^4 power law. Within the experimental error both fits agree reasonably well with the Rayleigh scattering model.

We also note that we are currently not aware of a theoretical lineshape analysis of transmission spectra of periodic nano-hole arrays. Our analysis indicates that the above lineshape model gives a surprisingly good agreement and we hope that these results will stimulate efforts towards this direction.

5.4.6 Bandgap formation

In their original Nature paper (170), Ebbesen analyze the SP bandstructure in such nano-hole arrays and say that there measurements "clearly demonstrate the presence of gaps with energies around 30 to 50 meV". In their later article (174) the authors point out that "the presence of Wood's anomaly prevents the measurements of such gaps".

In figure 5.19, we show the resonance energies of the SP transmission resonances and Wood's anomalies as derived from the analysis of the angle-resolved transmission spectrum. Even if we account for the presence of the Woods' anomalies in the lineshape model, we clearly observe a splitting of about $\Delta E_g = 35$ meV between the resonance energies of SM[-1,0] (0.79 eV) and SM[1,0] (0.825 eV) at $k_x = 0$. The experimental data give clear evidence that this splitting is not due to the presence of the Wood's anomaly. It reflects the bandgap formation in these nanostructures.

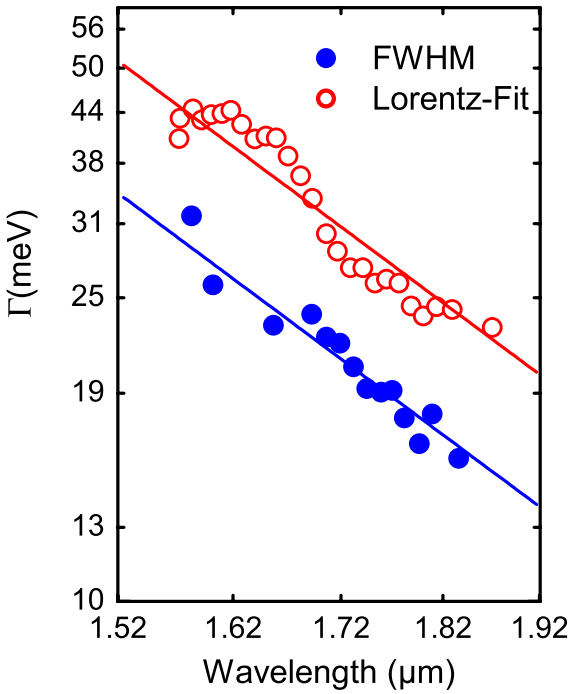

Figure 5.18: (filled circles) FWHM of the angle-dependent transmission spectrum. (open circles) Linewidth γ as extracted from the lineshape analysis of the angle-dependent transmission spectra. (solid lines) Fit to $\Gamma \propto \lambda^{-4}$. All data are plotted on a double logarithmic scale.

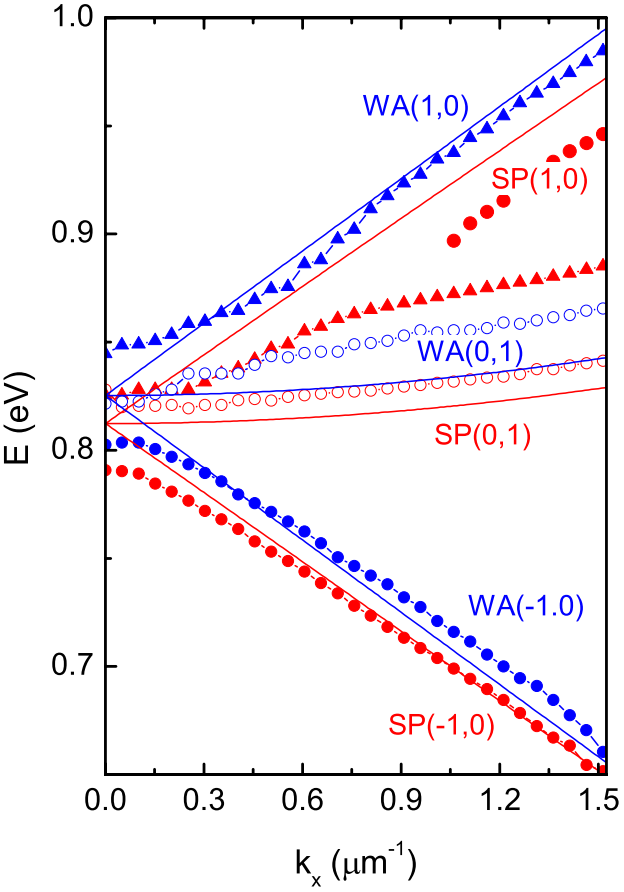

Figure 5.19: Dispersion relation of SP resonances and Wood's anomaly as extracted from the lineshape analysis of the angle-dependent transmission spectra.

The experimental data also show the SM[0,1] SP transmission resonance and allow to derive its dispersion relation. The resonance corresponds to grating diffraction of the incident light in the y-direction, orthogonal to the incident polarization. Its dispersion curve lies in between the SM[-1,0] and SM[1,0] SP resonances.

Formation of a full bandgap, characterized by the absence of enhanced transmission regardless of the k-vector of the incident light, requires that the linewidths γ of the SP resonances are smaller than the band gap ΔE_g. In the investigated sample, with a relative large hole radius of 150 nm, γ (30 meV) and ΔE_g (35 meV)are of similar magnitude and hence a full band gap is not formed. Yet, as shown in Fig. 5.17, variation of the hole radius allows to vary the lineshape over a wide range. These experiments indicate that it should indeed be possible to demonstrate full band gap formation in samples with smaller hole radii.

5.4.7 Hole size dependence of the SP scattering

Conclusion about macroscopic origin of the SP dumping stated in subsection 5.4.4 is fully supported by the hole size dependence of the SP scattering cross section. A row of 8 nano-holes, each separated by 10 μm, is placed at a distance of 50 μm from a periodic nano-hole array. Similar to Fig. 5.12(a), SP waves are generated in the array, propagate over 50 μm across the flat metal and are scattered at the single nano-holes. In the NF, interference patterns of incident and scattered SPs are formed and detected. At a tip-to-sample distance of 0.6 μm, however, these evanescent SP modes are not observed and only the photons scattered into the FF are collected. We find a drastic decrease in emission intensity by a factor of 30, as r is decreased from 267 to 110 nm [Fig. 5.20(a)]. The double logarithmic plot

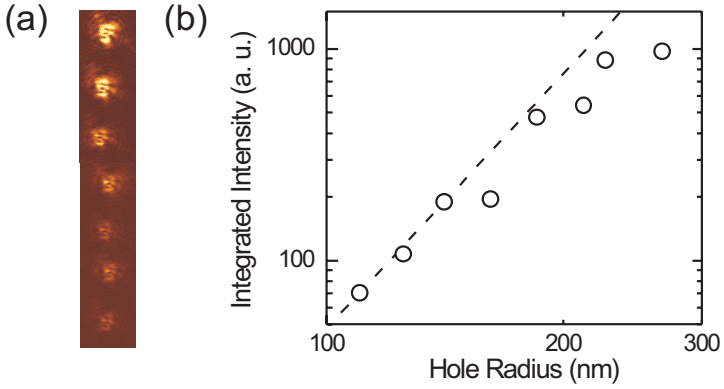

Figure 5.20: (a) FF emission patterns from various single holes with radii decreasing from 267 nm (top) to 110 nm (bottom). The images are recorded at a tip-to-sample distance of 0.6 μm. (b) Log-log plot of the spatially integrated emission intensity around the single holes plotted against their hole radius r. The dotted line is a fit to the fourth power of r.

in Fig. 5.20(b) shows for small radii indeed the predicted r^4 dependence. For the large radii a weaker dependence is observed, most likely reflecting that here the Rayleigh limit is no more applicable. The results show unambiguously that the damping time T_2 of SP excitations in these nano-hole arrays is limited by the finite SP lifetime $T_1 \simeq T_2/2$ due to Rayleigh scattering into FF radiation. Pure dephasing processes, reflecting phase fluctuations without population decay and Ohmic losses due to nonradiative SP decay into phonons are of minor importance. Hence, the SP lifetime is substantially shorter than on high-conductivity metal surfaces, where nonradiative losses may dominate (194; 192). The radiative damping mechanism in nano-hole arrays is similar to that in metal nanoparticles (193), yet here the damping times are typically an order of magnitude shorter, giving rise to linewidths of several hundreds of meV. This reduced damping is relevant for applications.

5.5 Superposition of polarization controlled surface waves in the near-field

5.5.1 Polarization dependence of NF emission patterns

In this section, we consider NF emission patterns in plasmonic crystals. It was shown in section 5.3 that the dominant contribution to the emission can be from the metal surfaces between the holes. When the polarization is along the symmetry direction, the NF emission pattern is dominated by stripes separated by lattice constant a_0, that run perpendicular to the polarization direction. While the stripes remain as the dominant feature, the detailed pattern varies critically depending on the excitation wavelength near the AM [1,0] or [0,1] SP resonances. This strongly suggests coherent superposition of many SP waves with various diffraction orders. In contrast, at the SM [1,1] SP resonance, localized light emission around the holes are seen. Comparison of reflection and transmission spectra reveals the inherent difference in the excitation and emission processes for these two modes.

We employ a NSOM (42; 45) (see section 3.3) in the transmission geometry. Details of experiment and the sample used are described in section 5.4.1.

Fig. 5.21(a) shows an atomic force microscopy image of our gold sample, while Fig. 5.21 (b), (c) and (d) show the NF emission pattern for horizontal, diagonal, and vertical polarizations respectively, which are indicated by the arrows. The wavelength of light is 780 nm, near the AM [1,0] or [0,1] resonance. While the transmission efficiency in the FF region is rather independent of polarization in these structures (176), the NF stripe-like patterns always appear in a perpendicular manner to polarization. As we also note in 5.3.2 the emission pattern is by no means localized around the holes. Instead, they are dominated by the stripe-like patterns, reflecting the symmetry of coherently propagating SP waves.

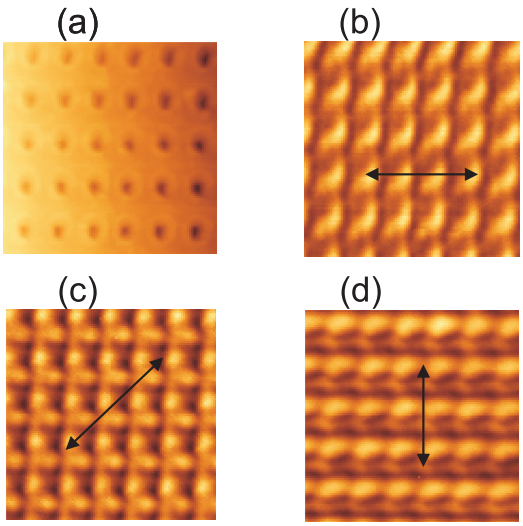

Figure 5.21: (a) topography of the gold sample with 760 nm period (b), (c), (d) near-field images at different polarizations (arrows). The scanning area is roughly 5 μm $\times$ 5 μm.

5.5.2 Wavelength dependence of NF emission patterns

We now discuss the wavelength dependence of the emission patterns on a different gold sample with period of 650 nm and hole size of 150 nm. Fig. 5.22(a) shows the FF transmission spectrum in the region of the AM [1,0] or [0,1] resonance. The arrows indicate the spectral positions: 750, 760, 780, and 800 nm at which the NF images Fig. 5.22(b), (c), (d), and (e) were taken, respectively. The polarization was vertical in all images. The images again show stripe-like patterns that run perpendicular to the polarization direction. The NF images are indeed dominated not by bright holes as our macroscopic expectation would dictate, but by these stripes. It also shows that the details of the images are very sensitive to the excitation wavelength. For instance, Fig. 5.22(b) shows two dominant stripes within one period along the vertical direction. On the other hand, Fig. 5.22 (c), (d), and (e) have one dominant stripe in one period, with localized emissions in between the stripes. Often, a fine structure exists within the localized emission, such as the dark spots in the middle of the localized emission as shown in Fig. 5.22(c).

The complicated, yet periodic emission patterns and also their strong wavelength dependence strongly suggest that patterns arise through the interference between coherently propagating SP's with different diffraction orders, characterized by the wavenumbers $\pm\frac{2\pi}{a_0}(m,n)$ for integers m and n. SP's propagate along the direction of polarization (vertical) which results in a relatively large excitation of the $m = 0$ modes while various lower m modes can be excited due to the scattering at holes during propagation. This explains the dominance of the stripe-like patterns that run perpendicular to the polarization direction.

In Fig. 5.22(d) and (e), we see localized emission between the strong stripe-like patterns. The natural question that arises then is whether these localized emission spots coincides with the position of holes. The location of emission spot at the center of holes requires constructive interference of each SP waves at exactly

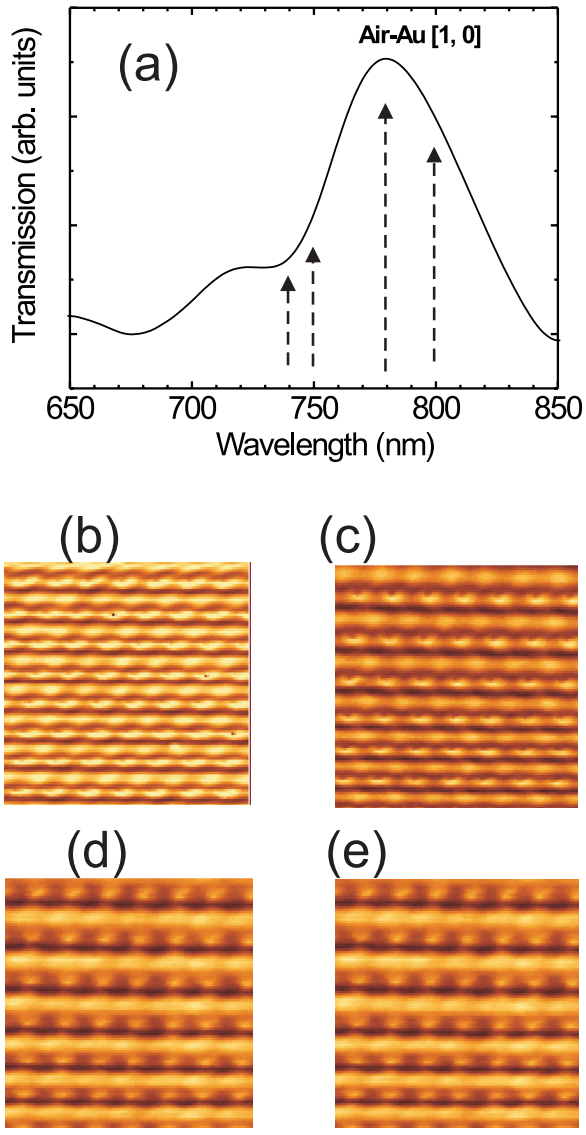

Figure 5.22: (a) Transmission spectrum of a gold sample with period=650 nm, hole size=150 nm, and film thickness=200 nm (b), (c), (d), and (e): NF emission patterns at wavelengths of 740, 750, 780, and 800 nm, respectively.

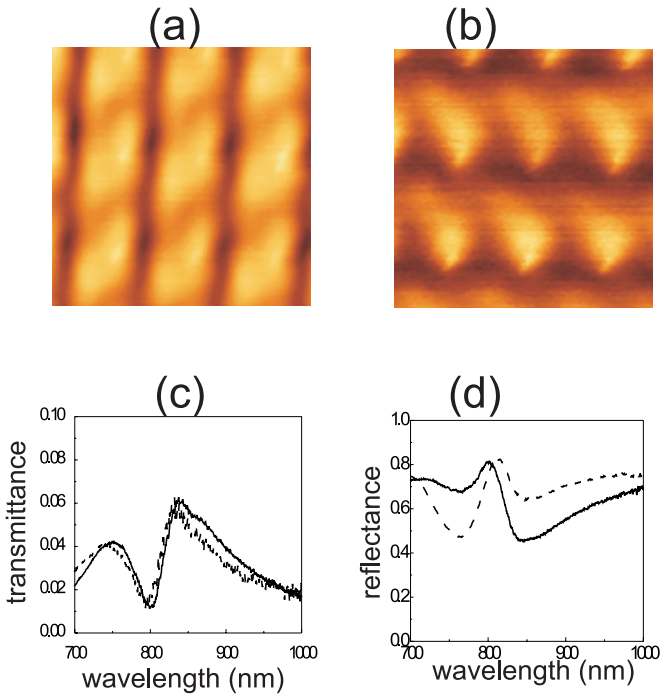

Figure 5.23: Comparison of air-silver [1,0] mode and sapphire-silver [-1,1] mode. Holes with 250 nm diameter were made on 200 nm thick silver films with a period of 600 nm. Emission patterns of the AM [1,0] mode using λ = 730 nm. (b) Emission patterns of the SM [1,1] resonance using λ= 830 nm. (c) FF transmission spectrum for the said silver sample. Dotted lines: transmission spectrum when the incident light is from the air side. (d) FF reflection spectrum taken from the sapphire side in the same case as (c). Dotted lines: reflection spectrum when the incident light is from the air side.

that position. This condition can easily be broken by minor inhomogeneities and asymmetries in the sample and the experimental conditions. Experimentally, the localized emission spot never exactly coincides with the position of holes.

5.5.3 AM vs. SM resonances

Thus far, we have concentrated on the nanoscopic emission patterns at the AM SP resonance when light impinges upon the SM interface. To compare emission patterns of both the AM and the SM resonances, additional experiments were performed using a silver sample with a shorter period. In Fig. 5.23(a), the NF emission pattern is shown for a silver sample with 600 nm period and 250 nm hole diameter, which is excited at the [1,0] or [0,1] AM SP resonance (750 nm). Again, the emission stripes are perpendicular to the polarization direction. However, at the [1,1] SM SP resonance (830 nm), the NF emission pattern (Fig. 5.23(b)) is completely different from those at the AM SP resonance. Bright regions are

located near the holes and elongated along the diagonal direction of lattice, which is the propagation direction of the SM [1,1] resonance. We observe elongation along the diagonal directions regardless of the incident polarization. The failure to form stripe patterns in this case can be understood from the fact that the SM [1,1] mode is excited on the opposite side from the measuring tip and transmitted through holes. Due to the difference in dielectric constants of sapphire and air, the wavelength of the SM SP wave, once exposed to the AM interface, is considerably longer and is completely incommensurate with the lattice constant. Hence the coherent propagating SP waves cannot be excited and the emission is mainly from the holes. Yet, it is very interesting that the localized emission patterns are aligned along the diagonal direction, which still reflects the symmetry of the [1,1] SM SP that is strongly excited at the other interface.

Our interpretation of the different emission patterns for the AM and the SM modes is largely corroborated by analyzing FF transmission and reflection spectra. Fig. 5.23(c) shows the transmission spectra for the same silver sample illuminated from either the sapphire side (solid lines) or the air side (dotted lines). Both spectra show clear maxima at both the AM and the SM resonances and their strength is more or less independent on the choice of illumination side.

While the two transmission peaks have similar peak strength, the reflection spectra shown in Fig. 5.23(b) display two dips of very different strengths. Furthermore, unlike the transmission spectrum peaks, the two reflection dips reverse their relative strengths when the illumination and the probing sides are interchanged. In both cases, the larger dip corresponds to the SP excitation at the illuminated interface side.

In the case of the sapphire-side illumination, the AM SP is excited mainly through the coupling provided by the nano-holes. This is why the reflection dip for the AM SP mode is much weaker than that for the SM mode. On the other hand, once excited at the AM side, AM SP can propagate, interfere, and can be converted to light very efficiently. This picture is supported when we normalize the transmission peak to the reflection dip. For the [1,0] or [0,1] AM resonance, the strength of the transmission peak (about 0.04 in Fig. 5.22c) divided by that for the reflection dip (about 0.1 in Fig. 5.23d) is approximately 40 %, highlighting the extremely high conversion efficiency. In contrast, the SM mode, with transmission peak at about 0.05 and the reflection dip at about 0.4 has much lower conversion efficiency, suggesting that most of its energies are spent on the SM interface due to heat loss. Once exposed to the AM interface, this mode emits light mostly around the holes.

5.6 Evolution of the near-field patterns into the far-field

5.6.1 Transition form near-field into far-field

In this section, we concern ourselves with the question as to how the near-field emission patterns of the periodic metal nano-hole array structures (170; 173) change in the far-field region. In the preceding section (5.5), we showed unambiguously that the interference between various modes of coherently propagating SP (171; 172; 144) waves dominate the NF emission patterns. This result is somewhat different from earlier works assuming that light gets scattered by the holes (176), and the field intensity reaches maxima near the center of the holes (182). The FF patterns evolved from the NF patterns have been studied (178; 179; 180; 185; 186) mostly from the theoretical point of view. The relationship between higher diffraction orders and the smaller penetration depth has been also identified (185).

We study the emission pattern of the transmitted light through a metal nano-hole array structure as a function of the distance z between the tip and the sample using the metal coated tip with a sub-100 nm resolution. The sample is a golden plasmonic crystall described in section 5.4.1. It should be noted that, because the SP resonance peak (AM [1,0] mode at 825 nm) is so broad, its high energy tail can easily reach to the wavelength that is equal to a_0 or even smaller. Thus the first-diffraction order SP, which becomes propagating for the excitation wavelength below a_0, get enhanced due to resonance. This results in a very efficient wave-guide revealing the interference between the zeroth and the first order SP diffraction orders, which in turn generates spatially sinusoidal waves. How far the sinusoidal wave can reach into the FF depends sensitively on the excitation wavelength relative to the lattice constant. For wavelengths smaller than a_0, sinusoidal patterns have been observed up to z=15 μm, with a very good peak-to-valley ratio of about 6.

Fig. 5.24(a) shows a NF image taken with an excitation wavelength at very close to the AM [1,0] resonance, with a horizontal polarization. The stripes perpendicular to the polarization direction can be seen. At z=0.2 μm, the signal intensity decreases while a clear sinusoidal pattern appears (Fig. 5.24(b)). The sinusoidal pattern becomes homogeneous with increasing z, so that in the FF region, only the homogeneous part remains. In other words, we truly reach the FF region at z=1.8μm where the intensity becomes constant (Fig. 5.24(c)). Shown in Fig. 5.24(d) are the horizontal cross-sectional scans of Fig. 5.24(a), (b), and (c) from left to right, respectively.

The fact that the emission pattern becomes a very good spatial sinusoidal function with a peak-to-valley ratio larger than 6 at z=0.2 μm implies that the sinusoidal function at z=0.2 μm results mainly from the interference between the zeroth order diffraction and the evanescent first order diffraction. Strong presence of any other modes would have made the pattern more complicated. At larger z,

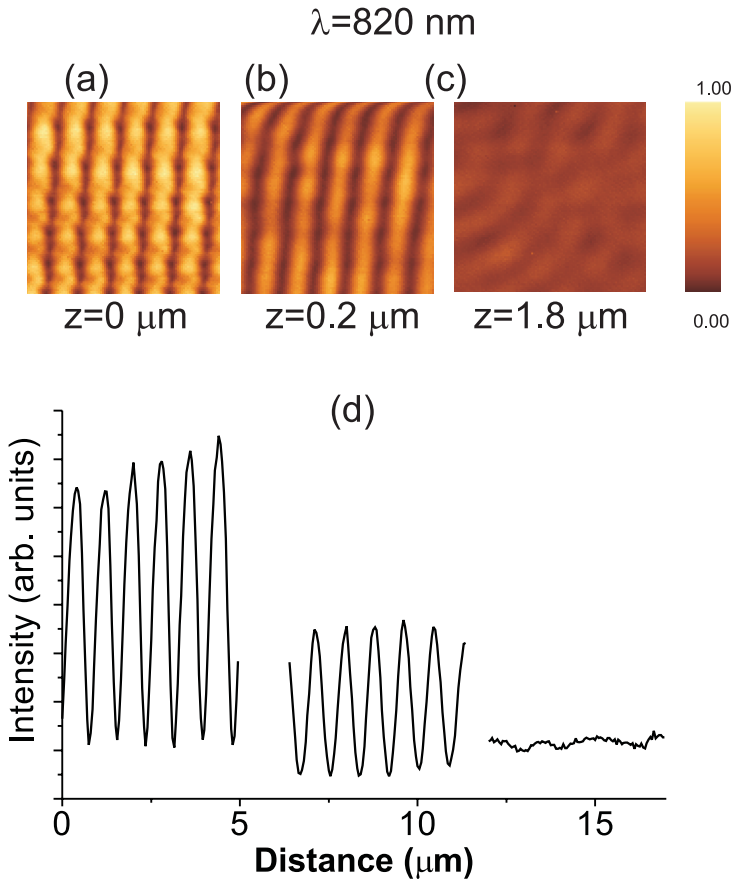

Figure 5.24: (a) NF emission pattern for λ = 820 nm (b) emission pattern at z=0.2 μm. (c) emission pattern for z=1.8 μm. (d) cross sectional scans for the images for z=0, 0.2 and 1.8 μm from left to right, respectively

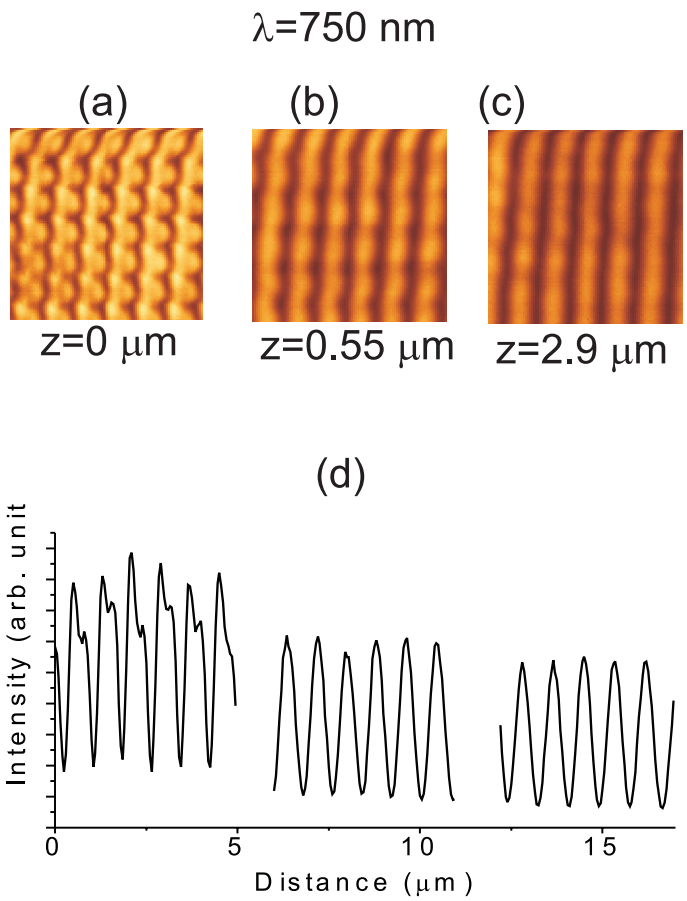

Figure 5.25: (a) NF emission pattern for λ = 750 nm (b) emission pattern at z=0.55 μm. (c) emission pattern for z=2.9 μm. (d) cross sectional scans for the images for z=0, 0.55 and 2.9 μm from left to right, respectively

only the homogeneous, zeroth order diffraction mode survives.

The situation is markedly different at the high energy side of the AM [1,0] resonance. Shown in Fig. 5.25(a), (b), and (c) are emission images at NF, at 0.55 and 2.9 μm, respectively, at the excitation wavelength of 750 nm < a_0. The initially complicated pattern becomes sinusoidal with increasing z but instead of becoming homogeneous, the sinusoidal pattern persists up to the FF, with an excellent peak-to-valley ratio (Fig. 5.25(d)). It is clear that the essential nature of the pattern evolution into the FF is very different depending on whether the wavelength is larger or smaller than the lattice constant. To investigate this issue further, we performed a detailed study on the wavelength dependence.

Shown in Fig. 5.26 (a) are the FF region images at z=2 μm, at various wavelengths around the lattice constant. At this distance, clear sine waves are seen only up to λ = 760nm and the stripe patterns become blurred beyond this point and at the excitation wavelength of 780 nm, the image becomes more or less homoge-

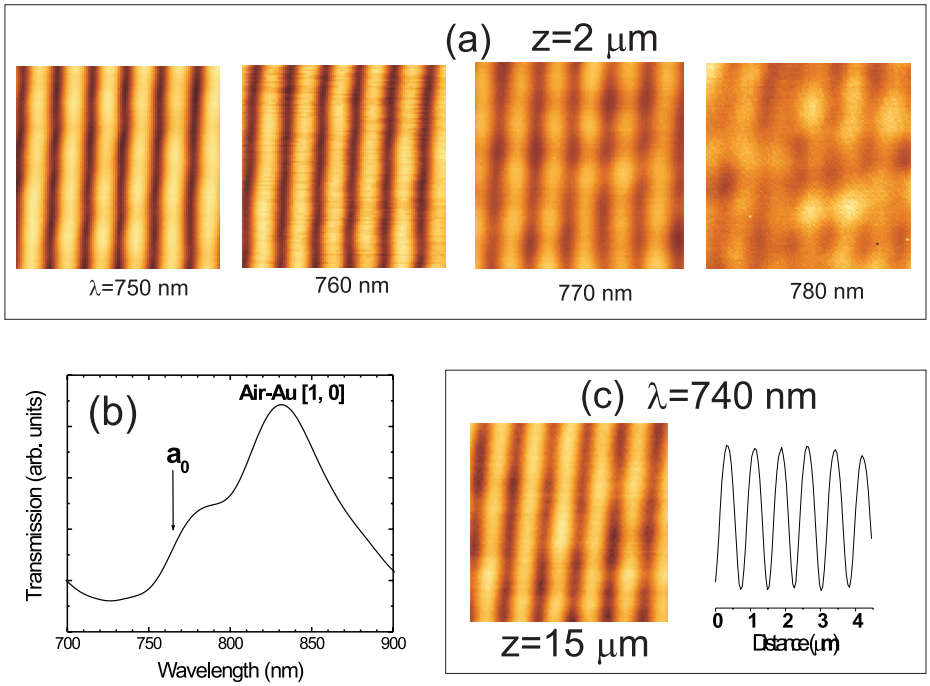

Figure 5.26: (a) Emission patterns at a FF region z=2 μm, with varying wavelengths: 750, 760, 770, and 780 nm from left to right, respectively. (b) Transmission spectrum of the sample (c) emission pattern and cross section images at λ=740 nm and z=15 μm.

neous. Fig. 5.26(b) shows the FF transmission spectrum at normal incidence, which clearly demonstrates that $\lambda = a_0$ is well within the broad resonance. Strikingly, at $\lambda = 740nm$, the spatial sine wave with an excellent peak-to-valley ratio persists up to z=15 μm, while the spot size is about 30 μm.

5.6.2 Nano-slits diffraction analysis

To understand these experimental results, we first observe that radiation patterns in the FF region appear as stripes perpendicular to the polarization direction without the details of hole geometry. This suggests that the experiment can be described effectively by the diffraction at the nano-slits instead of a rigorous diffraction analysis at the nano-holes. Thus in the following analysis, we assume that the system is symmetric along the slit direction which we choose to be the y-direction and the electric field lies in the *xz*-plane. In the experimental measurement, we are concerned with the component of Poynting vector, $S_z = E_x H_y$, along the propagation

direction. The electric field component E_x is given by

$$\begin{aligned} E_x &= \sum_{n=0}^{r-1} 2A_n \cos(2\pi n x/a)(\cos k_n z \cos wt + \sin k_n z \sin wt) \\ &+ \sum_{n=r}^{\infty} 2A_n \exp(-h_n z)\cos(2\pi n x/a)\cos wt \end{aligned} \tag{5.44}$$

where the coefficients A_n specify relative excitation strengths of each modes and the integer r denotes the number of radiating modes determined such that the n-th order momenta k_n and h_n defined by

$$\begin{aligned} k_n &\equiv \sqrt{(2\pi/\lambda)^2 - (2\pi n/a)^2}, n < r \\ h_n &\equiv \sqrt{-(2\pi/\lambda)^2 + (2\pi n/a)^2}, n \geq r \end{aligned} \tag{5.45}$$

are all real. The other nonvanishing electric field component E_z and the magnetic field component H_y can be obtained directly by integrating the Maxwell's equations

$$\begin{aligned} \nabla \cdot \vec{E} &= \partial_x E_x + \partial_z E_z = 0 \\ -\frac{1}{c}\frac{\partial H_y}{\partial t} &= \partial_z E_x - \partial_x E_z. \end{aligned} \tag{5.46}$$

We first assume that the excitation wavelength is shorter than the lattice constant $\lambda \lesssim a_0$ so that $r = 2$. That is, only the homogeneous ($n = 0$) and the first order diffraction mode ($n = 1$) are radiating. Higher order modes ($n \geq 2$) are all evanescent. In the far field region where only the radiating mode survives, the time average of S_z can be readily computed to give

$$\begin{aligned} \bar{S_z} &= 2A_0^2 + 2A_1^2\Big(1 - \frac{\lambda^2}{a^2}\Big)^{-1/2}\cos^2(2\pi x/a) \\ &+ 2A_0A_1\Big(1 + \Big(1 - \frac{\lambda^2}{a^2}\Big)^{-1/2}\Big)\cos(2\pi x/a)\cos\Big(\frac{2\pi z}{\lambda}\Big[\Big(1 - \frac{\lambda^2}{a^2}\Big)^{1/2} - 1\Big]\Big) \end{aligned} \tag{5.47}$$

More realistically, we could include a Gaussian modulation factor to E_x along the x-direction in order to simulate the Gaussian beam and derive the time average of S_z following the same procedure as above. Fig 5.27 displays analytic results computed this way. In all cases, lattice constant was 750 nm and the Gaussian beam width was assumed to be 8 μm.

Shown in Fig. 5.27(a) and (b) are the time average of S_z and its transversal cross sections at two different z values for $\lambda < a_0$. The zeroth and the first order modes persist into the FF region and their mutual interference generates well defined sinusoidal waves. In particular, Eq. (5.47) shows that the period of patterns in the propagation direction is given by

$$\lambda\Big[1 - \Big(1 - \frac{\lambda^2}{a^2}\Big)^{1/2}\Big)\Big]^{-1} \tag{5.48}$$

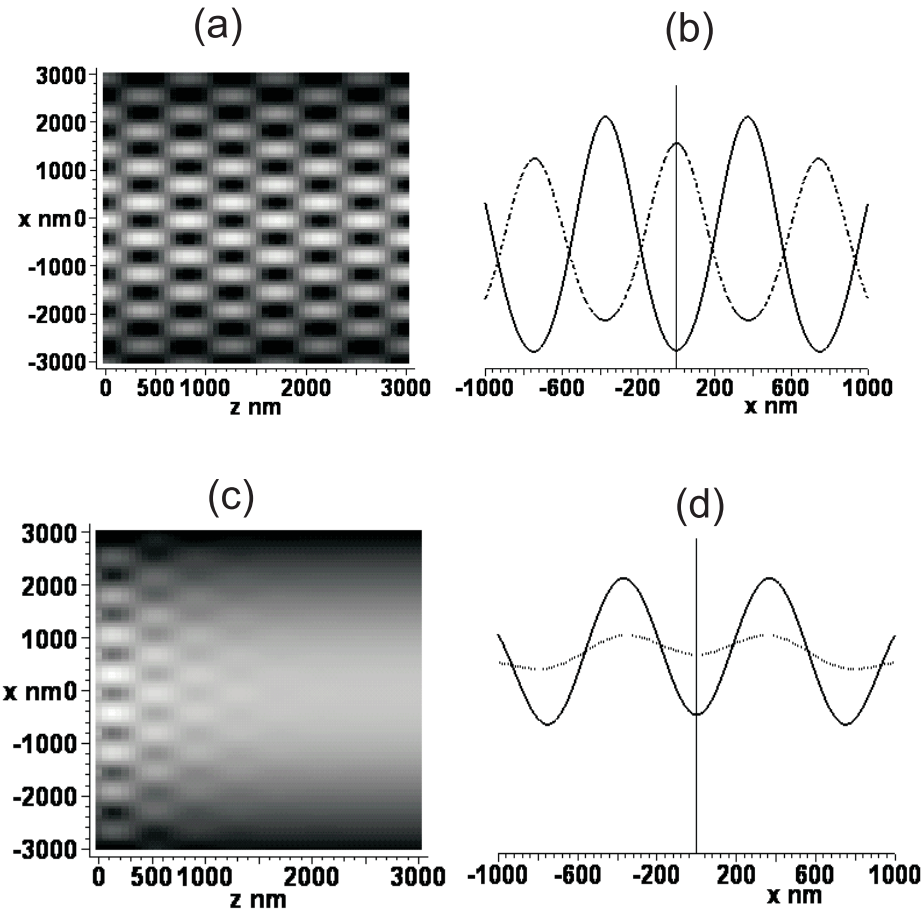

Figure 5.27: Emission patterns predicted by a diffraction theory. (a) time averaged Poynting vector component $\bar{S}_z$ of a Gaussian beam and (b) its transversan cross section for $\lambda = 740nm < a_0 = 750nm$, for z=500 (solid line) and 1000 nm (dashed lines). (c) and (d) are for the case $\lambda = 780nm > a_0 = 750nm$. z values for (d) are 200 (solid line) and 1000 nm (dashed lines).

which is bigger than the wavelength λ.

If the excitation wavelength is longer than the lattice constant $\lambda > a_0$, only the homogeneous ($n = 0$) mode radiates. All other diffraction modes are evanescent and decay exponentially. Figs. 5.27(c) and 5.27(d) show the time average of S_z and the cross sections in this case where the interference between the homogeneous and the decaying diffraction modes can be explicitly observed.

These results agree well with a direct numerical result using the 2-dimensional FDTD method (see section 5.2.4) as shown in Fig. 5.28. Fig. 5.28a describes the

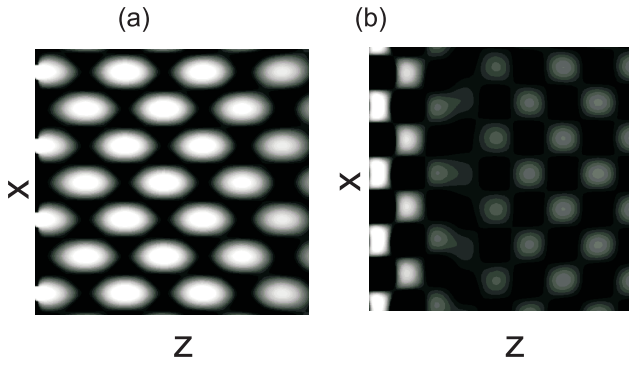

Figure 5.28: FDTD calculation of $\bar{S_z}$ showing emission patterns for the cases (a) $\lambda = 740nm < a_0 = 750nm$ and (b) $\lambda = 780nm > a_0 = 750nm$.

case for $\lambda < a_0$, where sinusoidal patterns in the time averaged Poynting vector component are apparent. When $\lambda > a_0$, the pattern becomes much more homogeneous in qualitative agreement with experiments, but the sinusoidal pattern does not completely die out with increasing z.

Chapter 6

Conclusions

The thesis is devoted to near-field spectroscopy of semiconductor device structures and plasmonic crystals. The two types of systems considered here are examples of the current state of nanoscale technology. While first type of devices already finds broad applications in scientific and production utilities, the second type is still paving its way towards the realization of its potential. Both systems impose a lot of difficulties on experimentalist since the characteristic dimensions lay well below resolution limit of conventional optics. The reliable characterization of such systems is a challenge because of the requirement that the technique is able to overcome diffraction limit. Here, NSOM is instrument of choice as it break through Abbe limit by using evanescent component of light and so providing necessary resolution on order of few tenth of nanometers.

The main results of investigation can be summarized as follows:

- We have demonstrated the potential of NSOM based optical spectroscopy for non-destructive optoelectronic device analysis and particularly for investigation of InGaAs/
  AlGaAs high-power lasers. This technique provides direct visualization of the effect of the waveguide design on the number of guided modes and their spatial profile.

- We have straightforward and separate access to the properties of both optically active segments of novel monolithically stacked lasers with two nominally identical waveguide segments separated by a tunnel junction. We have recorded and consistently interpreted data on device electroluminescence and lasing, as well as photoluminescence and photocurrent. We find that reduced laser emission from one laser segment is caused by a larger trap concentration withing or in the vicinity of the quantum wells of this laser segment. In the unbiased devices the potential gradient in the segment with reduced laser emission is significantly larger than in the other.

- The interplay between surface recombination and diffusion creates a spatial carrier profile at the cleaved interface between a quantum-well and air.

Analysis of this profile requires experimental data with sub-diffusion length spatial resolution as well as solution of a two-dimensional balance equation. Our NSOM data as well as numerical simulation allow the independent determination of surface recombination velocity and the diffusion length. We find that the surface recombination velocity decreases with increasing intensity due to the saturation of nonradiative defect states. In this transition region the overall shape of the luminescence signal depends strongly on the excitation intensity.

- We have directly resolved surface plasmon excitations on the nonilluminated metal surface of periodic nano-hole arrays in metal films. The evolution of the near-field into the far-field emission is studied and conditions are shown where an essentially spatially sinusoidal far-field pattern can generated.

- We have shown that interference of coherently propagating SP waves generate symmetric emission patterns dominated by stripe-like patterns that are perpendicular to the polarization direction and therefore the propagation direction of SP. We explained different emission patterns for the AM and the SM SP modes. Our work shows that the metal nano-structures can be a very efficient source of coherently propagating surface waves that can find its applications in nano-optic devices.

- We have experimentally studied the damping of SP excitations in metallic nano-hole arrays. By relating nanoscopic SP propagation, ultrafast light transmission and optical spectra, we demonstrate that the transmission spectra of these plasmonic band gap structures are homogeneously broadened. The spectral line shape and damping times are dominated by Rayleigh scattering of SP into light and can varied over a wide range by controlling the resonance energy and/or hole radius. This opens the way towards designing SP nano-optic devices and spatially and spectrally tailoring light-matter interactions on nanometer length scales.

- We have studied, both experimentally and theoretically, the evolution of the NF emission pattern into the FF region in a metal film punctured with periodic arrays of nano-holes. The complicated NF pattern becomes spatially sinusoidal or homogeneous depending on whether the wavelength is larger or smaller than the lattice constant.

The experiments presented in this thesis justify near-field microscopy as a sensitive nanoscopic tool for nondestructive analysis of optoelectronic and plasmonic devices. The scientific results of our investigation throw light on an interesting new aspect of the origin of enhanced transmission and gives new insight into the microscopic physics of light-matter-interaction that take place in diode lasers and plasmonic nanostructures. The knowledge gained in the scope of this work provides important new information on how to tailor the optical properties

of optoelectronic devices and plasmonic nanostructures for use in future applications.

Appendix A

First order perturbation theory

The accurate solution of a physical problem can be found in a relatively small number of simplest cases. Most problems lead to a very complicated sets of equations that can not be solved exactly. Often such problems include values of different orders. Many of them can be relatively small. It can happen that by neglecting them the problem becomes simplified to a level that a make is possible to be solve exactly. In this case a first step towards a solution is finding an accurate solution of this simplified problem, and second–the approximate calculation of the corrections that appear from the small parameters neglected in the simplified approach. The general method for the calculation of such corrections is the *perturbation theory*.

Perturbation theory finds highly successful application in many fields of modern physics. Most of them are in quantum mechanics which formalism we will use in this subsection (7).

Perturbation theory can be applied to physical systems which Hamiltonian can be presented in the form

$$H = H_0 + W, \tag{A.1}$$

where W is the small correction (*perturbation*) to the "unperturbed" operator H_0. Solutions for the unperturbed system

$$H_0 \left|\varphi_i^{(0)}\right\rangle = E^{(0)} \left|\varphi_i^{(0)}\right\rangle \tag{A.2}$$

are known. Here values of $|\varphi_p^{(0)}\rangle$ are the set for the orthonormal basis:

$$\begin{aligned} \langle\varphi_i \left|\varphi_{i'}\right\rangle &= \delta_{ii'}, \\ \sum_i |\varphi_i\rangle\langle\varphi_i| &= 1. \end{aligned} \tag{A.3}$$

We can write $W = \lambda\hat{W}$ where λ is a dimesionless unit that is much smaller than unity. So we can consider the Hamilton operator as dependent on the perturbation parameter λ:

$$H(\lambda) = H_0 + \lambda\hat{W}. \tag{A.4}$$

For $\lambda = 0$ we get the unperturbed Hamilton operator H_0.

In order to solve the perturbation problem we need to find a solution for

$$H(\lambda)\,|\psi(\lambda)\rangle = E(\lambda)\,|\psi(\lambda)\rangle. \tag{A.5}$$

We presume that $E(\lambda)$ and $|\varphi(\lambda)\rangle$ can be expanded into powers of λ:

$$\begin{aligned} E(\lambda) &= \varepsilon_0 + \lambda\varepsilon_1 + \ldots + \lambda^q\varepsilon_q + \ldots, \\ |\psi(\lambda)\rangle &= |0\rangle + \lambda\,|1\rangle + \cdots + \lambda^q\,|q\rangle + \ldots \end{aligned} \tag{A.6}$$

We connect this expansion together with definition (A.4) of $H(\lambda)$ into equation (A.5),

$$\Big(H_0 + \lambda\,\hat{W}\Big)\Big[\sum_{q=0}^{\infty}\lambda^q\,|q\rangle\Big] = \Big[\sum_{q'=0}^{\infty}\lambda^{q'}\varepsilon_{q'}\Big]\Big[\sum_{q=0}^{\infty}\lambda^q\,|q\rangle\Big], \tag{A.7}$$

and demand, that this equation will be satisfied for any small λ. The collection of coefficients bring us

– for 0^{th}-order therms in λ

$$H_0\,|0\rangle = \varepsilon_0\,|0\rangle, \tag{A.8}$$

–for 1^{st}-order therms

$$(H_0 - \varepsilon_0)\,|1\rangle + (\hat{W} - \varepsilon_1)\,|0\rangle = 0, \tag{A.9}$$

–for 2^{nd}-order therms

$$(H_0 - \varepsilon_0)\,|2\rangle + (\hat{W} - \varepsilon_1)\,|1\rangle - \varepsilon_2\,|0\rangle = 0, \tag{A.10}$$

–for qth-order therms

$$(H_0 - \varepsilon_0)\,|q\rangle + (\hat{W} - \varepsilon_1)\,|q-1\rangle - \varepsilon_2\,|q-2\rangle - \ldots - \varepsilon_q\,|0\rangle = 0. \tag{A.11}$$

We will concentrate on first three equations neglecting all terms that are higher than second order of λ. In this way we can find first order corrections.

We assume that vector $|\psi(\lambda)\rangle$ is normalized. We can choose the phase in the way that scalar product $\langle 0\,|\psi(\lambda)\rangle$ remain real. For the 0^{th} order, the vector $|0\rangle$ should be normalized,

$$\langle 0|0\rangle = 1. \tag{A.12}$$

The phase, however, is arbitrary. For the 1^{st} order we can write the square of the norm for $|\psi(\lambda)\rangle$ as

$$\begin{aligned} \langle\psi(\lambda)|\psi(\lambda)\rangle &= [\langle 0| + \lambda\langle 1|]\,[|0\rangle + \lambda\,|1\rangle] + O(\lambda^2) \\ &= \langle 0|0\rangle + \lambda\,[\langle 1|0\rangle + \langle 0|1\rangle] + O(\lambda^2) \end{aligned} \tag{A.13}$$

It is clear from equation (A.12) that this expression in first order is equal to 1 when terms with λ disappear. We choose the phase in the way that scalar product $\langle 0|1\rangle$ is real (λ is real). That means

$$\langle 0|1\rangle = \langle 1|0\rangle = 0. \tag{A.14}$$

Analogous considerations for second order of λ result in

$$\langle 0|2\rangle = \langle 2|0\rangle = -\frac{1}{2}\langle 1|1\rangle \tag{A.15}$$

and for qth order

$$\begin{aligned}\langle 0|q\rangle &= \langle q|0\rangle \\ &= -\frac{1}{2}[\langle q-1|1\rangle + \langle q-2|2\rangle + \ldots + \langle 2|q-2\rangle + \langle 1|q-1\rangle].\end{aligned} \tag{A.16}$$

Equation (A.8) expresses that $|0\rangle$ is the eigenvector of H_0 with the eigenvalue ε_0 belonging to the spectrum of H_0. We consider one particular value of ε_0, i. e. one eigenvalue E_n^0 of the H_0. This value belongs to the particular eigenvector $|\varphi_n\rangle$. We want to investigate the changes to this undisturbed energy and the corresponding steady state due to adding of the perturbation W to the Hamilton operator.

For this we will use perturbation equations (A.8) – (A.11) with conditions (A.12), (A.14) – (A.16). The eigenvalue of $H(\lambda)$ is approaching E_n^0 when $\lambda \to 0$. So we can write

$$\varepsilon_0 = E_n^{(0)}. \tag{A.17}$$

Taking into account (A.8) it follows that $|0\rangle$ is proportional to $|\varphi_n\rangle$. Both vectors $|0\rangle$ and $|\varphi_n\rangle$ are normalized (see eq.(A.12)), and we can choose

$$|0\rangle = |\varphi_n^{(0)}\rangle. \tag{A.18}$$

For $\lambda \to 0$ we will get again the unperturbed steady state $|\varphi_n\rangle$ with the same phase.

Projecting the vector $\langle\phi_n|$ onto the equation (A.9) we get

$$\langle\varphi_n|(H_0 - \varepsilon_0)|1\rangle + \langle\varphi_n|(\hat{W} - \varepsilon_1)|0\rangle = 0. \tag{A.19}$$

Taking into account (A.3) and (A.19)

$$\varepsilon_1 = \langle\varphi_n|\hat{W}|0\rangle = \langle\varphi_n|\hat{W}|\varphi_n\rangle. \tag{A.20}$$

Now we can write energy expression with first order correction for perturbation $W = \lambda\hat{W}$:

$$E_n(\lambda) = E_n^0 + \langle\varphi_n|W|\varphi_n\rangle + O(\lambda^2). \tag{A.21}$$

The projection (A.19) uses not all the information from equation (A.9). We should project this equation on all vectors of the $|\varphi_p\rangle$-basis that are different from $|\varphi_n\rangle$.

$$\langle\varphi_p|(H_0 - E_n^0)|1\rangle + \langle\varphi_p|(\hat{W} - \varepsilon_1)|\varphi_n\rangle = 0 \qquad (p \neq n). \tag{A.22}$$

The term $\varepsilon_1\langle\varphi_p|\varphi_n\rangle$ is equal to zero because eigenvectors of H_0 are orthogonal for different eigenvalues. Equation (A.22) will be

$$(E_p^0 - E_n^0)\langle\varphi_p|1\rangle + \langle\varphi_p|\hat{W}|\varphi_n\rangle = 0, \tag{A.23}$$

and now we can write an expression for the coefficients from expansion of the vector $|1\rangle$ for all unperturbed basis states except $|\varphi_n\rangle$:

$$\langle\varphi_p|1\rangle = \frac{1}{E_n^0 - E_p^0}\langle\varphi_p|\hat{W}|\varphi_n\rangle = 0 \qquad (p \neq n). \tag{A.24}$$

The last coefficient $\langle\varphi_n|1\rangle$ is equal to zero because of condition (A.14) (from Eq. (A.18) it follows that $|\varphi_n\rangle$ is equal to $|0\rangle$),

$$\langle\varphi_n|1\rangle = 0. \tag{A.25}$$

Now we know the vector $|1\rangle$ can be expanded in the $|\varphi_p\rangle$-basis:

$$|1\rangle = \sum_{p\neq n} \frac{\langle\varphi_p|\hat{W}|\varphi_n\rangle}{E_n^0 - E_p^0}|\varphi_p\rangle. \tag{A.26}$$

Thus we can write the first order perturbation $W = \lambda\hat{W}$ for the unperturbed state $|\psi_n(\lambda)\rangle$ of the H:

$$|\psi_n(\lambda)\rangle = |\varphi_n\rangle + \sum_{p\neq n} \frac{\langle\varphi_p|W|\varphi_n\rangle}{E_n^0 - E_p^0}|\varphi_p\rangle + O(\lambda^2) \tag{A.27}$$

The first order correction for the state vector is the linear superposition of the all $|\varphi_n\rangle$ different unperturbed states. One can say, that perturbation W cause the mixture of the state $|\varphi_n\rangle$ with other eigenstates $|\varphi_p\rangle$ of the H_0.

Appendix B

Derivation of the dispersion relation of SPs on a surface of a semi-infinite solid

The layer system, Fig. B.1, see also Fig. 5.1, has an interface (1/2), e.g., metal

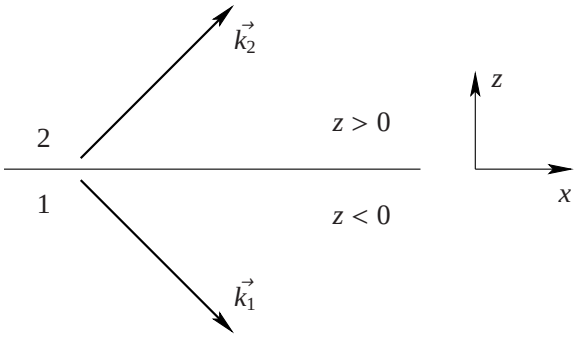

Figure B.1: Scheme of the interface between two media i and k on which the SP propagates. The wavevector k_i has the components k_x and k_{iz}, see (5.3) or (B.14).

(ε_1)/air (ε_2), on which a $p-$polarized wave propagates in the x direction. There is no y dependence. We describe the fields in the media (1) and (2) as follows:

$$\begin{aligned} z > 0 \quad \mathbf{H}_2 &= (0, H_{y2}, 0) \exp i(k_{x2}x + k_{z2}z - \omega t) \\ \mathbf{E}_2 &= (E_{x2}, 0, E_{z2}) \exp i(k_{x2}x + k_{z2}z - \omega t) \end{aligned} \tag{B.1}$$

$$\begin{aligned} z < 0 \quad \mathbf{H}_1 &= (0, H_{y1}, 0) \exp i(k_{x1}x + k_{z1}z - \omega t) \\ \mathbf{E}_1 &= (E_{x1}, 0, E_{z1}) \exp i(k_{x1}x + k_{z1}z - \omega t). \end{aligned} \tag{B.2}$$

These fields have to fulfill Maxwell's equations:

$$\operatorname{rot} \mathbf{H}_i = \varepsilon_i \frac{1}{c} \frac{\partial E_i}{\partial t} \tag{B.3}$$

$$\operatorname{rot} \mathbf{E}_i = -\frac{1}{c} \frac{\partial H_i}{\partial t} \tag{B.4}$$

$$\operatorname{div} \varepsilon_i \mathbf{E}_i = 0 \tag{B.5}$$

$$\operatorname{div} \mathbf{H}_i = 0, \tag{B.6}$$

together with the continuity relations

$$E_{x1} = E_{x2} \tag{B.7}$$

$$H_{y1} = H_{y2} \tag{B.8}$$

$$\varepsilon_1 E_{z1} = \varepsilon_2 E_{z2}. \tag{B.9}$$

From (B.8) follows the continuity of

$$k_{x1} = k_{x2} = k_x. \tag{B.10}$$

Equation (B.3) gives

$$\begin{aligned} \frac{\partial H_{yi}}{\partial z} &= -\varepsilon_i E_{xi} \frac{\omega}{c} \quad \text{or} \\ +k_{z1} H_{y1} &= +\frac{\omega}{c} \varepsilon_1 E_{x1} \\ +k_{z2} H_{y2} &= -\frac{\omega}{c} \varepsilon_2 E_{x2}. \end{aligned} \tag{B.11}$$

Equation (B.11) together with (B.7, B.8) yield

$$\begin{aligned} H_{y1} - H_{y2} &= 0 \\ \frac{k_{z1}}{\varepsilon_1} H_{y1} + \frac{k_{z2}}{\varepsilon_2} H_{y2} &= 0. \end{aligned} \tag{B.12}$$

To obtain a solution, the determinant D_0 has to be zero

$$D_0 = \frac{k_{z1}}{\varepsilon_1} + \frac{k_{z2}}{\varepsilon_2} = 0. \tag{B.13}$$

This is the dispersion relation of the SPs in the system. Further we get from (B.3, B.4, B.11)

$$k_x^2 + k_{zi}^2 = \varepsilon_i \left(\frac{\omega}{c} \right)^2. \tag{B.14}$$

From (B.13) together with (B.14) follows

$$k_x = \frac{\omega}{c} \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \tag{B.15}$$

Let us assume for the medium $\varepsilon_2 = 1$ (air) and for the medium ε_1 (metal) $\varepsilon_1 < 0$, $|\varepsilon_1| > 1$, then $k_x > \omega/c$ and k_{zi} becomes imaginary. The fields have their maxima at the surface plane $z = 0$ and decay in both z directions exponentially as is characteristic for surface waves.

Equations (B.13, B.14, B.15) here corresponds to equations (5.2, 5.3, 5.4).

Appendix C

Room temperature PL spectra from GaAs:C

This chapter devoted to experiment designed to clarify specific line shape of the PL spectrum from doped GaAs. It was demonstrated before that such PL spectrum demonstrate clear double peak structure (201; 202; 203). Different models were proposed to explain this phenomenon (201; 202; 203).

A Ti:sapphire laser system (see subsection 3.3.2) is used for interband excitation at a wavelength of 735 nm. Spectral selection is provided by dispersing the PL in a 0.25 m monochromator. The experiments are performed with a spot size of the laser focus of $\varnothing_{\mathrm{FWHM}} = 200\,\mu$m. Steady-state PL spectra are measured with standard equipment.

The room temperature NSOM describe in section 3.3 was employed in an illumination/collection geometry for PL measurements at cleaved sample edges (204).

In order to estimate the effects of doping on the DOS, Urbach parameters E_0, i.e., the slopes of the exponential part of the absorption constant below the energy gap E_g (205), are determined by photocurrent (206) and PL measurements. We find values of $E_0 = 16 - 24$ meV at 295 K and E_0 tend to increase with the doping level.

The investigated DH are grown by MOCVD (see subsection 4.1.1) in a horizontal Aixtron AIX200 reactor on GaAs substrates. Trimethylgallium [$Ga(CH_3)_3$] as a group-III precursor together with Arsin (AsH_3) enable intrinsic incorporation of carbon into the GaAs layers by choosing the V/III ratio near unity at a growth temperature below 600 C. The group of clad samples with $Ga_xIn_{1-x}P$ are investigated. The $Ga_xIn_{1-x}P$ barriers are grown at 580 °C with a mole fraction of $x = 0.51$ to achieve lattice matching with GaAs [cf. Fig. C.1].

Figures C.2 shows steady-state room-temperature PL spectra measured with time-integrating detection for cw excitation. Data from four samples clad by InGaP barriers and with carrier concentrations of $p_0 = (1 - 5) \times 10^{19}$ cm^{-3} are compared.

All PL spectra display a clear double peak structure that exhibits a systematic

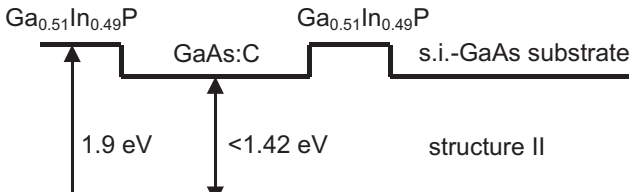

Figure C.1: Schematic of the heterostructures under investigation: carbon doped GaAs layers clad between

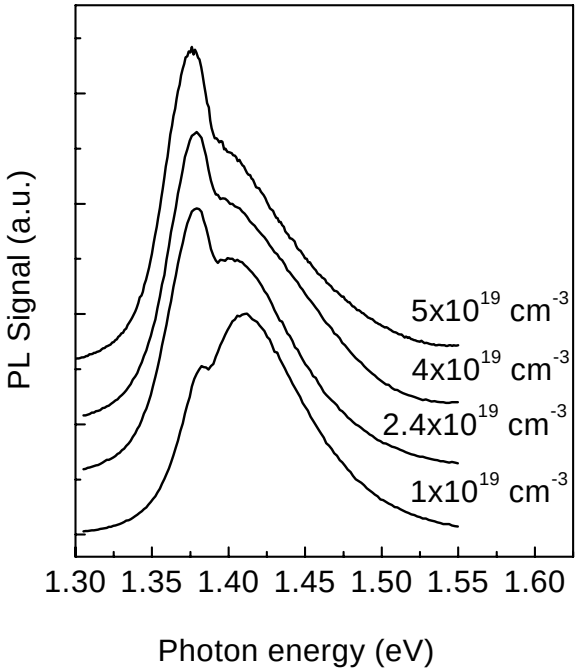

Figure C.2: Room-temperature cw PL spectra for four samples clad by GaInP barriers with carbon doping concentrations of 1, 2.4, 4, and 5×10^{19} cm^{-3} and a thickness of the GaAs:C layer of 1 μm. The excitation wavelength is 735 nm.

line shape variation with increasing p_0. The systematic dependence on doping concentration led the authors of Ref. (201) to conclude that the lower energy peak is a defect-related transition. On the other hand, it was suggested that the double structure arises from two different contributions to the PL: (202; 203) (i) The "regular" PL emitted into the direction of the collection optics and (ii) PL that is emitted into the substrate direction, propagates through the substrate, and is back reflected at the GaAs/Air interface at the bottom of the substrate. These two contributions are schematically illustrated in Fig. C.3(a). For the back-reflected PL contribution, the substrate acts as a spectral edge filter strongly suppressing luminescence above the substrates band gap energy. Within this model the observed systematic increase of the lower energy PL contribution to the double peak structure with doping is caused by the redshift of the PL emission due to band gap renormalization within the heavily doped layer that is of major importance in highly doped materials (202).

A clear distinction between the two alternative explanations can be made by comparing macroscopic spectra to those recorded with a NSOM. Figure C.3(c) compares a conventional cw PL spectrum obtained in FF PL geometry [see Fig. C.3(a)] and a cw spectrum that was excited and collected through the fiber tip of

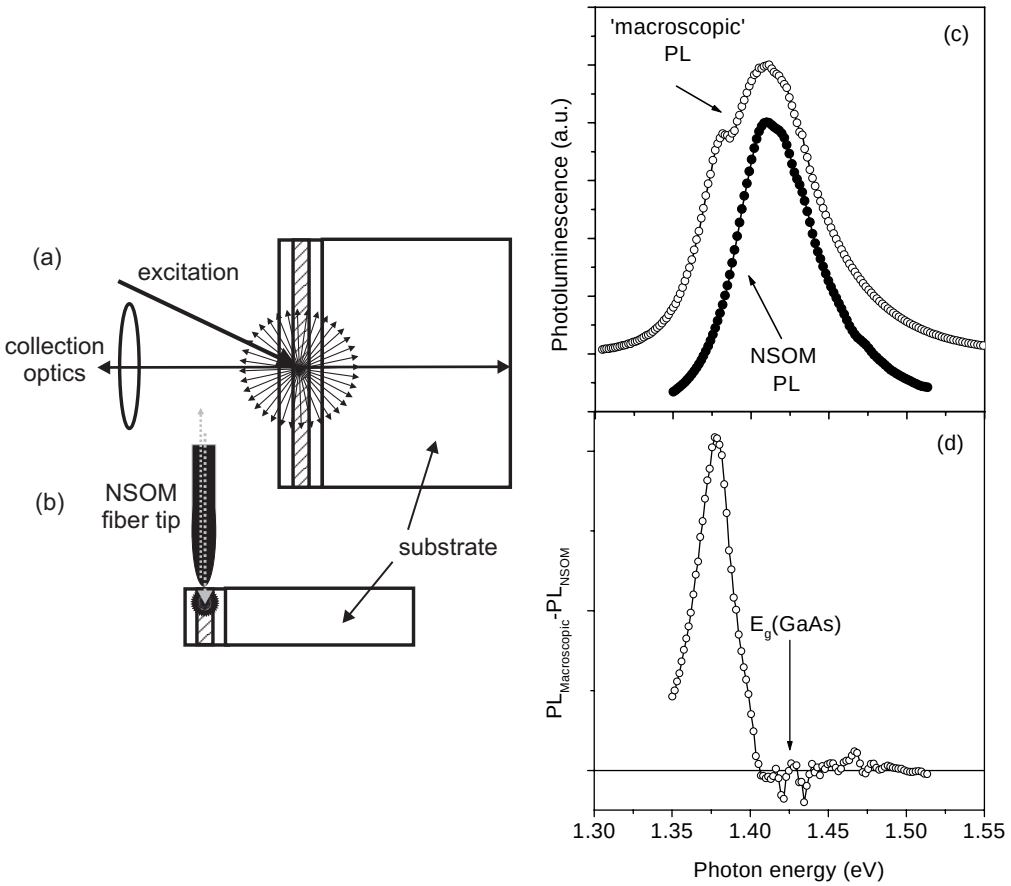

Figure C.3: (a) Schematic of the macroscopic cw PL excitation and detection scheme. The PL data are obtained by excitation through the top cladding by creating an excitation spot of $\varnothing = 200\ \mu$m. (b) Schematic of the NSOM based cw PL excitation and detection scheme. The NSOM is operated in illumination/collection geometry, i.e., the sample is locally excited through the NSOM fiber and PL is collected through the same fiber. (c) Macroscopic and NSOM based cw PL data from the same sample ($p_0 = 1 \times 10^{19}$ cm^{-3}). The macroscopic PL spectrum shows a clear double peak structure, whereas a single PL emission peak is observed in the NSOM configuration. (d) The difference spectrum between normalized macroscopic and NF spectra shows a clear cutoff slightly below the band edge of the substrate material. This assigns the second, low energy peak in the macroscopic PL to emission that propagates through the substrate and is back reflected from the GaAs/Air interface at the bottom of the substrate.

the NSOM as schematically shown in Fig. C.3(b). In this microscopic geometry back-reflected PL contributions propagating through the substrate and being picked up by the NSOM tip again are negligible and we expect to collect only PL being directly emitted in the direction of the NSOM fiber probe. Indeed, we observe a single-peaked PL spectrum with a substantially reduced FWHM of about 62 meV. The difference between normalized macroscopic and NF spectra [Fig. C.3(d)] shows a clear cutoff slightly below the band edge position of the substrate. This directly demonstrates that the second lower energy peak in the FF spectrum can be unequivocally assigned to PL that is emitted into the substrate direction, propagates through the substrate, and is back reflected at the bottom of the sample. The line broadening of the regular PL peak seen in the NF spectrum is close to the value of $1.8 \times kT = 46$ meV expected for room-temperature k-conserving spontaneous emission without degeneracy and $E_0 = 0$. The remaining broadening is easily explained by the presence of defect related tail states described by an Urbach parameter of $E_0 = 16 - 24$ meV.

References

[1] J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (John Wiley & Sons, Inc., 1998).

[2] J. D. Joannopoulos, R. D. Meade, and J. N. Winn, *Photonic Crystals: Molding the Flow of Light* (Princeton University Press, 1995).

[3] E. Yablonovitch, T. J. Gmitter, and K. M. Leung, Phys. Rev. Lett. **67,** 2295–2298 (1991).

[4] E. Yablonovitch, T. J. Gmitter, R. D. Meade, A. M. Rappe, K. D. Brommer, and J. D. Joannopoulos, Phys. Rev. Lett. **67,** 3380–3383 (1991).

[5] C. Kittel, *Introduction to Solid State Physics*, seventh ed. (John Wiley & Sons, Inc., New York, 1996).

[6] N. W. Ashcroft and N. D. Mermin, *Silid State Physics* (Saunders College Publishing, West Washington Square, Philadelphia, PA 19105, 1988).

[7] C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantenmechanik* (Walter de Gruyter GmbH & Co. KG, D-10785 Berlin, 1999), in 2 Teilen, Zweite Auflage.

[8] M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*, fourth ed. (Pergamon Press Ltd., Headington Hill Hall, Oxford, 4 & 5 Fitzroy Square, London W.1, 1970).

[9] E. Hecht, *Optics*, fourth ed. (Addison Wesley Longman, Inc., Reading, Massachusetts, 2002).

[10] B. D. Terris, H. J. Mamin, D. Rugar, W. R. Studenmund, and G. S. Kino, Appl. Phys. Lett. **65,** 388 (1994).

[11] M. Vollmer, H. Giessen, W. Stolz, L. Ghislain, and V. Elings, Appl. Phys. Lett. **74,** 1791 (1999).

[12] Q. Wu, R. D. Grober, D. Gammon, and D. S. Katzer, Phys. Rev. Lett. **83,** 2652 (1999).

[13] J. Christen, M. Grundmann, and D. Bimberg, J. Vac. Sci. and Technol. B **9,** 2358 (1991).

[14] C. A. Warwick, E. Kapon, D. M. Hwang, E. Colas, and L. Nunes, Phys. Rev. B **45,** 6333 (1992).

[15] A. Gustafsson, L. Samuelson, J. O. Malm, G. Vermeire, and P. Demmeester, Appl. Phys. Lett. **64,** 695 (1995).

[16] L. Abraham, A. Veider, C. Schönenberger, H. P. Meier, D. J. Arent, and S. F. Alvarado, Appl. Phys. Lett. **56,** 1564 (1990).

[17] D. Gammon, E. S. Snow, and D. S. Katzer, Appl. Phys. Lett. **67,** 2391 (1995).

[18] R. J. Warburton, C. Schäflein, D. Haft, F. Bickel, A. Lorke, K. Karrai, J. M. Garcia, W. Schönfeld, and P. Petroff, Nature **405,** 926 (2000).

[19] J. J. Finley, A. D. Ashmore, A. Lemaître, D. J. Mowbray, M. S. Skolnick, I. E. Itskevich, P. A. Maksym, M. Hopkinson, and T. F. Krauss, Phys. Rev. B **63,** 3307 (2001).

[20] J.-Y. Marzin, J.-M. Gérard, A. Izraël, D. Barrier, and G. Bastard, Phys. Rev. Lett. **73,** 716 (1994).

[21] A. Kuther, M. Bayer, A. Forchel, A. Gorbunov, V. B. Timofeev, F. Schäfer, and J. P. Reithmaier, Phys. Rev. B **58,** R7508 (1998).

[22] M. Bayer, O. Stern, P. Hawrylak, S. Fafard, and A. Forchel, Nature **405,** 923 (2000).

[23] J. Heltzer, M. Wegener, G. Khitrova, and H. M. Gibbs, In *Proceedings of QELS 2001*, (Baltimore, 2001).

[24] M. A. Paesler and P. J. Moyer, *Near-Field Optics: Theory, Instrumentation, and Applications* (John Wiley & Sons, Inc., John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, 1996).

[25] S. I. Bozhevolnyi, "Near-Field Optics of Nanostructured Surfaces," in *Optics of Nanostructured Materials, Wiley Series in Lasers and Applications*, V. A. Markel and T. F. George, eds., (John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, 2001), Chap. Three, pp. 73–142.

[26] A. J. den Dekker and A. van den Bos, J. Opt. Soc. Am. A **14,** 547 (1997).

[27] E. Betzig and J. K. Trautman, Science **257,** 189 (1992).

[28] E. H. Synge, Philos. Mag. **6,** 356 (1928).

[29] D. Courjon, K. Sarayeddine, and M. Spajer, Opt. Commun. **71,** 23 (1989).

[30] R. C. Reddick, R. J. Warmack, and T. L. Ferrel, Phys. Rev. B **39,** 767 (1989).

[31] F. de Fornel, J. P. Goudonnet, L. Salomon, and E. Lesniewska, In *Proc. SPIE*, **1139,** 77 (1989).

[32] D. V. Labeke and D. Barchiesi, J. Opt. Soc. Am. A **10,** 2193 (1993).

[33] F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, Appl. Phys. Lett. **65,** 1623 (1994).

[34] N. H. P. Moers, R. G. Tack, N. F. van Hulst, and B. Bölger, J. Appl. Phys. **75,** 1254 (1994).

[35] Y. Inouye and S. Kawata, Opt. Lett. **19,** 159 (1994).

[36] U. C. Fisher, J. Koglin, and H. Fuchs, J. Microsc. **176,** 231 (1994).

[37] R. Bachelot, P. Gleyzes, and A. C. Boccara, Opt. Lett. **20,** 1924 (1995).

[38] W. Denk and D. W. Pohl, J. Vac. Sci. Technol. B **9,** 510 (1991).

[39] O. J. F. Martin and C. Girard, Appl. Phys. Lett. **70,** 705 (1997).

[40] L. Aigouy, A. Lahrech, S. Grésillon, H. Cory, A. C. Boccara, and J. C. Rivoal, Opt. Lett. **24,** 187 (1999).

[41] J. A. O'Keefe, J. Opt. Soc. Am. **46,** 359 (1956).

[42] D. W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44,** 651 (1984).

[43] A. Harootunian, E. Betzig, M. Isaacson, and A. Lewis, Appl. Phys. Lett. **49,** 674 (1986).

[44] E. Betzig, M. Isaacson, and A. Lewis, Appl. Phys. Lett. **51,** 2088 (1987).

[45] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, Science **251,** 1468 (1991).

[46] S. Jiang, H. Ohsawa, K. Jamada, T. Pangaribuan, M. Othsu, K. Imai, and A. Ikai, Jpn. J. Appl. Phys. **31,** 2282 (1992).

[47] G. A. Valaskovic, M. Holton, and G. M. Morrision, Appl. Opt. **34,** 1215 (1995).

[48] H. Heinzelmann and D. W. Pohl, Appl. Phys. A **59,** 89 (1994).

[49] A. G. T. Ruiter, M. H. P. Moers, A. Jalocha, and N. F. van Hulst, Ultramicroscopy **61,** 139 (1995).

[50] S. Madsen, N. C. R. Holme, P. S. Ramanujam, S. Hvilsted, J. M. Hvam, and S. J. Smith, Ultramicroscopy **71,** 65 (1998).

[51] T. Saiki, S. Mononobe, M. Ohtsu, N. Saito, and J. Kusamo, Appl. Phys. Lett. **68,** 2612 (1996).

[52] D. I. Kavaldjiev, R. Toledocrow, and M. Vaeziravani, Appl. Phys. Lett. 67 (1995).

[53] A. H. Larosa, B. I. Yacobson, and H. D. Hallen, Appl. Phys. Lett. **67,** 2597 (1995).

[54] C. Lienau, A. Richter, and T. Elsaesser, Appl. Phys. Lett. **69,** 325 (1996).

[55] M. Stahelin, M. A. Bopp, G. Tarrach, A. J. Meixner, and I. Zschokkegranacher, Appl. Phys. Lett. **68,** 2603 (1996).

[56] E. Betzig, J. K. Trautman, J. S. Weiner, T. D. Harris, and R. Wolfe, Appl. Opt. **31,** 4563 (1992).

[57] A. Jalocha, M. H. P. Moers, A. G. T. Ruiter, and N. F. van Hulst, Ultramicroscopy **61,** 221 (1995).

[58] T. Lacoste, T. Huser, R. Prioli, and H. Heinzelmann, Ultramicroscopy **71,** 333 (1998).

[59] C. W. Hollars and R. C. Dunn, Rev. Sci. Instrum. **69,** 1747 (1998).

[60] M. Muranishi, K. Sato, S. Hosaka, A. Kikukawa, T. Shintani, and K. Ito, Jpn. J. Appl. Phys. **36,** L942 (1997).

[61] J. A. Veerman, A. M. Otter, L. Kuipers, and N. F. van Hulst, Appl. Phys. Lett. **72,** 3115 (1998).

[62] S. J. Bukofsky and R. D. Grober, App. Phys. Lett. **71,** 2749 (1997).

[63] P. Hoffmann, B. Dutoit, and R.-P. Salath, Ultramicroscopy **61,** 165 (1995).

[64] M. N. Islam, X. K. Zhao, A. A. Said, S. S. Mickel, and C. F. Vail, Appl. Phys. Lett. **71,** 2886 (1997).

[65] D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi, Appl. Phys. Lett. **68,** 2491 (1996).

[66] P. Lambelet, A. Sayah, M. Pfeffer, C. Psilipona, and Marquis-Weible, Appl. Opt. **37,** 7289 (1998).

[67] R. Stöckle, C. Fokas, V. Deckert, R. Zenobi, B. Sick, B. Hecht, and U. P. Wild, Appl. Phys. Lett. **75,** 160 (1999).

[68] U. Dürig, D. W. Pohl, and F. Rohner, J. Appl. Phys. **59,** 3318 (1986).

[69] W. E. Moerner, T. Plakthotnik, T. Irngartinger, U. P. Wild, D. W. Pohl, and B. Hecht, Phys. Rev. Lett. **73,** 2764 (1994).

[70] G. Guttroff, J. M. Keto, C. K. Shih, D. Anselm, and B. G. Streetman, Appl. Phys. Lett. **68,** 3620 (1996).

[71] D. W. Pohl, U. C. Fischer, and U. Dürig, In *Proc. SPIE*, **897,** 84 (1998).

[72] E. Betzig, P. L. Finn, and J. S. Weiner, Appl. Phys. Lett. **60,** 2484 (1992).

[73] R. Toledo-Crow, P. C. Yang, Y. Chen, and M. Vaez-Iravani, Appl. Phys. Lett. **60,** 2957 (1992).

[74] F. F. Froehlich and T. D. Milster, Appl. Phys. Lett. **70,** 1500 (1997).

[75] C. Durkan and I. V. Shvets, J. Appl. Phys. **79,** 1219 (1996).

[76] C. Durkan and I. V. Shvets, J. Appl. Phys. **80,** 5659 (1996).

[77] A. V. Zvyagin, J. D. White, M. Kourogi, M. Kozuma, and M. Ohtsu, Appl. Phys. Lett. **71,** 2541 (1997).

[78] M. J. G. P. G. Blome, J. Schöfer, and R. G. Ulbrich, Appl. Phys. Lett. **68,** 307 (1996).

[79] P. J. Moyer, S. Kammer, K. Walzer, and M. Hietschold, Ultramicroscopy **61,** 291 (1995).

[80] R. D. Grober, T. D. Harris, J. K. Trautman, and E. Betzig, Rev. Sci. Instrum. **65,** 626 (1994).

[81] C. Chun, P.-K. Wie, and W. Fann, Appl. Phys. Lett. **67,** 3835 (1995).

[82] K. Karrai and R. D. Grober, Appl. Phys. Lett. **66,** 1842 (1995).

[83] F. Krausz, C. Spielmann, T. Brabec, E. Wintner, and A. J. Schmidt, Opt. Lett. **17,** 204 (1992).

[84] S. H. Ashworth, M. Joschko, M. Woerner, E. Riedle, and T. Elsaesser, Opt. Lett. **20,** 2120 (1995).

[85] R. D. Dupuis, IEEE J. Quantum. Electron. **23,** 651 (1987).

[86] *High-Power Diode Lasers: Fundamentals, Technology, Applications*, Vol. 78 of *Topics in Applied Physics*, R. Diehl, ed., (Springer, 2000).

[87] R. N. Hall, G. E. Fenner, J. D. Kingsley, T. J. Soltys, and R. O. Carlson, Phys. Rev. Lett. **9,** 366 (1962).

[88] N. H. Jr. and S. F. Bevacqua, Appl. Phys. Lett. **1,** 82 (1962).

[89] G. H. Olsen, C. J. Nuese, and M. Ettenberg, IEEE J. Quantum. Electron. **15,** 688 (1979).

[90] G. B. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practice* (Academic, San Diego, 1989).

[91] M. Razeghi, *The MOCVD Challenge, 1: A Survey of GaInAsP – InP for Photonic and Electronic Applications* (Institute of Physics, Bristol, 1989).

[92] M. Razeghi, *The MOCVD Challenge, 2: A Survey of GaInAsP – GaAs for Photonic and Electronic Device Applications* (Institute of Physics, Bristol, 1995).

[93] A. C. Jones and P. O'Brien, *CVD of Compound Semiconductors: Precursor Synthesis, Development and Applications* (VCH, Weinheim, 1997).

[94] J. S. Roberts and N. J. Mason, J. Cryst. Growth **68,** 422 (1984).

[95] S. D. Hersee, M. Krakowski, R. Blondeau, M. Baldy, B. D. Créoux, and J. P. Duchemin, J. Cryst. Growth **68,** 282 (1984).

[96] B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (John Wiley & Sons, Inc., 1991).

[97] L. A. Coldren and S. W. Corzine, *Diode Lasers and Photonic Integrated Circuits* (John Wiley & Sons, Inc., 1995).

[98] D. Marcuse, *Theory of Dielectric Optical Waveguides* (Academic, London, 1974).

[99] A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, London, 1983).

[100] A. G. Choo, H. E. Jackson, U. Thiel, G. N. de Brabander, and J. T. Boyd, Appl. Phys. Lett. **65,** 947 (1994).

[101] S. Z. Bourzeix, J. M. Moison, F. Mignard, F. Barthe, A. C. Boccara, C. Licoppe, B. Mersali, M. Allovon, and A. Bruno, **73,** 1035 (1998).

[102] E. J. Sanchez, L. Novotny, and X. S. Xie, Phys. Rev. Lett. **82,** 4014 (1999).

[103] C. Lienau, A. Richter, A. Klehr, and T. Elsaesser, Appl. Phys. Lett. **69,** 2471 (1996).

[104] I. Horsch, R. Kusche, O. Marti, B. Weigl, and K. J. Ebeling, J. Appl. Phys. (1996).

[105] S. K. Buratto, J. W. P. Hsu, E. Betzig, J. K. Trautman, R. B. Bylsma, C. C. Bahr, and M. J. Cardillo, Appl. Phys. Lett. **65,** 2654 (1994).

[106] M. S. Unlu, B. B. Goldberg, W. D. Herzog, D. Sun, and E. Towe, Appl. Phys. Lett. **67,** 1862 (1995).

[107] A. Richter, J. Tomm, C. Lienau, and J. Luft, Appl. Phys. Lett. **69,** 3981 (1996).

[108] J. W. Tomm, A. Jaeger, A. Bärwolff, T. Elsaesser, A. Gerhard, and J. Donecker, Appl. Phys. Lett. **71,** 2233 (1997).

[109] J. W. Tomm, T. Günther, C. Lienau, A. Gerhardt, and J. Donecker, J. Cryst. Growth **210,** 296 (2000).

[110] R. D. Grober, T. Rutherford, and T. D. Harris, Appl. Opt. **35,** 3488 (1996).

[111] K. J. Ebeling, *Integrated Optoelectronics* (Springer, Berlin, 1993).

[112] J. P. van der Ziel and W. T. Tsang, Appl. Phys. Lett. **41,** 499 (1982).

[113] J. C. Garcia, E. Rosencher, P. Collot, N. Laurent, J. L. Guyaux, B. Vinter, and J. Nagle, Appl. Phys. Lett. **71,** 3752 (1997).

[114] S. G. Patterson, G. S. Petrich, R. J. Ram, and L. A. Kolodziejski, Electron Lett. **35,** 395 (1999).

[115] C. Hanke, L. Korte, B. D. Acklin, M. Behringer, G. Herrmann, J. Luft, B. D. Odorico, M. Marchiano, and J. Wilhelmi, In *Proc. SPIE*, **3947,** 50 (2000).

[116] http://www.osram-os.com/news/news_power.html.

[117] V. Malyarchuck, J. W. Tomm, T. Günther, R. Müller, R. Kunkel, C. Lienau, and J. Luft, In *Proc. SPIE*, **4287,** 111 (2001).

[118] C. Lienau and T. Elsaesser, in *Semiconductors and Semimetals*, K.-T. Tsen, ed., (Academic Press, 2001), Vol. 67, pp. 39–108.

[119] T. Guenther, V. Malyarchuk, J. W. Tomm, R. Müller, and C. Lienau, Appl. Phys. Lett. **78,** 1463 (2001).

[120] V. Emiliani, F. Intonti, C. Lienau, T. Elsaesser, R. Nötzel, and K. H. Ploog, Phys. Rev. B **64,** 1553161 (2001).

[121] D. E. Aspnes, Surf. Sci. **132,** 406 (1983).

[122] R. Schatz and C. G. Bethea, J. Appl. Phys. **76,** 2509 (1994).

[123] L. K. Luke, J. Appl. Phys. **90,** 3413 (2001).

[124] H. Hillmer, A. Forchel, T. Kuhn, G. Mahler, and H. P. Meier, Phys. Rev. B **43,** 13992 (1991).

[125] K. Tai, T. R. Hayes, S. L. McCall, and W. T. Tsang, Appl. Phys. Lett. **53,** 302 (1988).

[126] H.-C. Ostendorf and A. L. Endrös, Appl. Phys. Lett. **71,** 3275 (1997).

[127] S. Krawczyk, M. Bejar, A. Khoukh, R. Blanchet, B. Sermage, D. Cui, and D. Pavlidis, J. Appl. Phys. **38,** 992 (1999).

[128] A. J. Sabbah and D. M. Riffe, J. Appl. Phys. **88,** 6954 (2000).

[129] J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures*, *Springer Series in Solid State Physics*, 2nd ed. (Springer, Berlin, 1999).

[130] H. Ichikawa, K. Inoshita, and T. Baba, Appl. Phys. Lett. **78,** 2119 (2001).

[131] F. Intonti, V. Emiliani, C. Lienau, T. Elsaesser, V. Savona, E. Runge, R. Zimmermann, R. Nötzel, and K. H. Ploog, Phys. Rev. Lett. **87,** 076801 (2001).

[132] R. Müller and C. Lienau, Appl. Phys. Lett. **76,** 3367 (2000).

[133] A. Richter, G. B. M. Süptitz, C. Lienau, T. Elsaesser, M. Ramsteiner, R. Nötzel, and K. H. Ploog, Phys. Rev. Lett. **79,** 2145 (1997).

[134] R. Müller and C. Lienau, J. Microsc. **202,** 339 (2001).

[135] A. P. Ongstad, D. J. Gallant, and G. C. Dente, Appl. Phys. Lett. **66,** 2730 (1995).

[136] R. K. Ahrenkiel, R. Ellingson, S. Johnston, and M. Wanlass, Appl. Phys. Lett. **72,** 3470 (1998).

[137] http://www.ioffe.rssi.ru/SVA/NSM/Semicond/GaInAs/.

[138] http://gams.nist.gov/serve.cgi/Module/FISHPACK/SEPELI/7993/.

[139] http://gams.nist.gov/serve.cgi/Package/FISHPACK/.

[140] H. B. Keller, *Numerical methods for two-point boundary-value problems* (Blaisdel, Waltham, Mass., 1968).

[141] P. Swarztrauber and R. Sweet, *Efficient fortran subprograms for the soulution of elliptic partial differential equations*, NCAR technical note, 1975, NCAR-TN/IA-109, pp. 135-137.

[142] J. W. Tomm, A. Bärwolff, T. Elsaesser, and J. Luft, Appl. Phys. Lett. **77,** 747 (2000).

[143] F. Intonti, V. Emiliani, C. Lienau, T. Elsaesser, R. Nötzel, and K. H. Ploog, Phys. Rev. B **63,** 075313 (2001).

[144] H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, 1988).

[145] R. H. Ritchie, Phys. Rev. **106,** 874 (1957).

[146] C. J. Powell and J. B. Swan, Phys. Rev. **118,** 640 (1960).

[147] W. L. Barnes, T. W. Preist, S. C. Kitson, and J. R. Sambles, Phys. Rev. B **54,** 6227 (1996).

[148] K. M. Ho, C. T. Chan, C. M. Soukoulis, R. Biswas, and M. Sigalas, Solid State Comm. **89,** 413 (1994).

[149] S. Fan, P. R. Villeneuve, R. D. Meade, and J. D. Joannopoulos, Appl. Phys. Lett. **65,** 1466 (1994).

[150] C. C. Cheng and A. Scherer, J. Vac. Sci. Technol. B **13,** 2696 (1995).

[151] S. G. Romanov, N. P. Johnson, A. V. Fokin, V. Y. Butko, H. M. Yates, M. E. Pemble, and C. M. S. Torres, Appl. Phys. Lett. **70,** 2091 (1997).

[152] G. Feiertag *et al.,* Appl. Phys. Lett. **71,** 1441 (1997).

[153] S. Y. Lin, J. G. Fleming, D. L. Hetherington, K. M. H. B. K. Smith, R. Biswas, S. R. K. M. M. Sigalas, W. Zubrzycki, and J. Bur, Nature **394,** 251 (1998).

[154] R. D. Meade, A. Devenyi, J. D. Joannopoulos, O. L. Alerhand, D. A. Smith, and K. Kash, J. Appl. Phys. **75,** 4753 (1994).

[155] P. L. Gourley, J. R. Wendt, G. A. Vawter, T. M. Brennan, and B. E. Hammons, Appl. Phys. Lett. **64,** 687 (1994).

[156] B. D'Urso, O. Painter, J. O'Brien, T. Tombrello, A. Yariv, and A. Scherer, J. Opt. Soc. Am. B **15,** 1155 (1998).

[157] M. Kanskar, P. Paddon, V. Pacradouni, R. Morin, A. Busch, J. F. Young, S. R. Johnson, J. MacKenzie, and T. Tiedje, Appl. Phys. Lett. **70,** 1438 (1997).

[158] S. G. Johnson and J. D. Joannopoulos, *Photonic Crystals: The Road from Theory to Practice* (Kluwer Academic Publishers, Boston, Dordrecht, London, 2002).

[159] M. M. Sigalas, K.-M. Ho, R. Biswas, and C. M. Soukoulis, "Photonic Crystals," in *Optics of Nanostructured Materials, Wiley Series in Lasers and Applications*, V. A. Markel and T. F. George, eds., (John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, 2001), Chap. One, pp. 1–38.

[160] S. Satpathy, Z. Zhang, and M. R. Salehpour, Phys. Rev. Lett. **64,** 1239 (1990).

[161] J. D. Joannopoulos, P. R. Villeneuve, and S. Fan, Nature **386,** 143 (1997).

[162] R. Biswas, C. T. Chan, M. Sigalas, C. M. Soukoulis, and K. M. Ho, "Photonic band gap materials," in *Photonic band gap materials*, C. M. Soukoulis, ed., (Kluwer, Dordrecht, The Netherlands, 1996), p. 23.

[163] J. B. Pendry and A. MacKinnon, Phys. Rev. Lett. **69,** 2772 (1992).

[164] J. B. Pendry, J. Mod. Opt. **41,** 209 (1994).

[165] K. S. Yee, IEEE Trans. Antennas Propagation **14,** 302 (1966).

[166] *The Finite Difference Time Domain Method for Electromagnetics*, K. Kunz and R. Luebbers, eds., (CRC Press, Boca Raton, FL, 1993).

[167] *Computational Electromagnetics*, K. Umashankar and A. Taflove, eds., (Artech House, Boston, MA, 1993).

[168] H. A. Bethe, Phys. Rev. **66,** 163 (1944).

[169] "C. J. Bouwkamp," Philips Res. Rep. **5,** 321 (1950).

[170] W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, Nature **391,** 667 (1998).

[171] G. Hass, M. H. Francombe, and R. W. Hoffman, *Physics of Thin Films* (Academic, New York, 1977), Vol. 9, p. 145.

[172] V. M. Agranovich and D. L. Mills, *Surface Polaritons* (North-Holland, Amsterdam, 1982).

[173] J. R. Sambles, Nature **391,** 641 (1998).

[174] H. F. Ghaemi, T. Thio, D. E. Grupp, T. W. Ebbesen, and H. J. Lezec, Phys. Rev. B **58,** 6779 (1998).

[175] T. J. Kim, T. Thio, T. W. Ebbesen, D. E. Grupp, and H. J. Lezec, Opt. Lett. **24,** 256 (1999).

[176] T. Thio, H. F. Ghaemi, H. J. Lezec, P. A. Wolff, , and T. W. Ebbesen, J. Opt. Soc. Am. B **16,** 1743 (1999).

[177] D. E. Grupp, H. J. Lezec, T. W. Ebbesen, K. M. Pellerin, and T. Thio, Appl. Phys. Lett. **77,** 1569 (2000).

[178] U. Schroter and D. Heitmann, Phys. Rev. B **58,** 15419 (1998).

[179] U. Schroter and D. Heitmann, Phys. Rev. B **60,** 4992 (1999).

[180] J. A. Porto, F. J. Garcia-Vidal, and J. B. Pendry, Phys. Rev. Lett. **83,** 2845 (1999).

[181] W.-C. Tan, T. W. Preist, and R. J. Sambles, Phys. Rev. B **62,** 11134 (2000).

[182] L. Salomon, F. Grillot, A. Zayats, and F. de Fornel, Phys. Rev. Lett. **86,** 1110 (2001).

[183] L. Martin-Moreno, F. J. Garcia-Vidal, H. J. Lezec, K. M. Pellerin, T. Thio, J. B. Pendry, and T. W. Ebbesen, Phys. Rev. Lett. **86,** 1114 (2001).

[184] Y. Takakura, Phys. Rev. Lett. **86,** 5601 (2001).

[185] S. J. McNab, R. J. Blaikie, and M. M. Alkaisi, J. Vac. Sci. Technol. B **18,** 2900 (2000).

[186] M. M. Alkaisi, R. J. Blaikie, S. J. McNab, R. Cheung, and D. R. S. Cumming, Appl. Phys. Lett. **75,** 3560 (1999).

[187] Q. Cao and P. Lalanne, Phys. Rev. Lett. **88,** 057403 (2002).

[188] S. I. Bozhevolnyi, J. Erland, K. Leosson, P. M. W. Skovgaard, and J. M. Hvam, Phys. Rev. Lett. **86,** 3008 (2001).

[189] M. M. J. Treacy, Appl. Phys. Lett. **75,** 606 (1999).

[190] C. Chicanne, T. David, R. Quidant, J. C. Weeber, Y. Lacroute, E. Bourillot, A. Dereux, G. C. des Francs, and C. Girard, Phys. Rev. Lett. **88,** 097402 (2002).

[191] S. C. Hohng *et al.*, Appl. Phys. Lett. **81,** 3239 (2002).

[192] B. Lamprecht, J. R. Krenn, G. Schider, H. Ditlbacher, M. Salerno, N. Felidj, A. Leitner, F. R. Aussenegg, and J. C. Weeber, Appl. Phys. Lett. **79,** 51 (2001).

[193] C. Sönnichsen, T. Franzl, T. Wilk, G. von Plessen, J. Feldmann, O. Wilson, and P. Mulvaney, Phys. Rev. Lett. **88,** 077402 (2002).

[194] M. van Exter and A. Lagendijk, Phys. Rev. Lett. **60,** 49 (1988).

[195] A. Dogariu, T. Thio, L. J. Wang, T. W. Ebbesen, and H. J. Lezec, Opt. Lett. **26,** 450 (2001).

[196] H. C. van de Hulst, *Light Scattering by Small Particles* (Dover Publications, New York, 1981).

[197] A. V. Shchegrov, I. V. Novikov, and A. A. Maradudin, Phys. Rev. Lett. **78,** 4269 (1997).

[198] U. Fano, J. Opt. Soc. Am. **31,** 213 (1941).

[199] A. Hessel and A. A. Oliner, Appl. Opt. **4,** 1275 (1965).

[200] M. Neviere, in *Electromagnetic theory of gratings*, Vol. 22 of *Springer Topics in Modern Physics*, R. Petit, ed., (Springer, Heidelberg, 1980).

[201] H. D. Chen, M. S. Feng, P. A. Chen, K. C. Lin, and J. W. Wu, Jpn. J. Appl. Phys., Part 1 **33,** 1920 (1994).

[202] J.-S. Lee, I. Kim, B.-D. Choe, W. G. Jeong, Y. K. Sin, and W. S. Min, J. Appl. Phys. **79,** 9278 (1996).

[203] Z. H. Lu and A. Majerfeld, J. Appl. Phys. **75,** 2648 (1994).

[204] C. Lienau, A. Richter, and J. W. Tomm, Appl. Phys. A: Mater. Sci. Process. **64,** 341 (1997).

[205] F. Urbach, Phys. Rev. **92,** 1324 (1953).

[206] J. W. Tomm, A. Bärwolff, A. Jaeger, T. Elsaesser, J. Bollmann, W. T. Masselink, A. Gerhardt, and J. Donecker, J. Appl. Phys. **84,** 1325 (1998).

Index